DOCUMENT RESUME

ED 300 648 CE 051 323

AUTHOR Zornitsky, Jeffrey; And Others
TITLE Establishing a Performance Management System for Targeted Welfare Programs. Research Report No. 88-14.
INSTITUTION Abt Associates, Inc., Washington, DC.
SPONS AGENCY National Commission for Employment Policy (DOL), Washington, D.C.
PUB DATE Aug 88
CONTRACT J-9-M-8-0012
NOTE 123p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS *Employment Programs; *Evaluation Criteria; Evaluation Methods; Management Systems; *Performance Factors; *Predictor Variables; Program Administration; Program Development; Program Evaluation; Standards; *Welfare Services
IDENTIFIERS *Aid to Families with Dependent Children; *Job Training Partnership Act 1982

ABSTRACT

A study examined a series of performance standards issues in the context of employment and training programs targeted toward the Aid to Families with Dependent Children (AFDC) population. Telephone interviews were conducted with representatives from the Job Training Partnership Act (JTPA) and departments of public welfare programs in nine states. These interviews were conducted to describe and document policies and procedures that encourage services for hard-to-employ individuals; existing welfare performance standards and adjustment procedures; key concerns with performance standards management; and major constraints in serving AFDC recipients, including the hard-to-employ. The second phase of the study entailed an assessment of the predictive validity of alternative performance measures that could be used for employment and training programs serving AFDC recipients. Fourteen of the more than 20 separate correlations evaluated were found to be significant predictors of earnings gains and reductions in welfare dependency. Average weekly earnings and average monthly benefit levels measured over the six months following termination were found to be the best predictors of both benchmarks. (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED300648

Establishing a Performance Management System for Targeted Welfare Programs

Jeffrey Zornitsky, Project Director
Mary Rubin

with the assistance of:
Dr. Stephen H. Bell
William Martin

Research Report No. 88-14

August 1988

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

A Study Prepared by Abt Associates, Inc. for the
National Commission for Employment Policy
1522 K Street N.W., Suite 300, Washington, D.C. 20005

BEST COPY AVAILABLE

CE 057 323

This report was prepared for the National Commission for Employment Policy under Contract No. J-9-M-8-0012. Since grantee/contractors conducting research and evaluation projects under government sponsorship are encouraged to express their own judgment freely, this report does not necessarily represent the official opinion or policy of the National Commission for Employment Policy. The authors are solely responsible for the contents of this report.

ACKNOWLEDGEMENTS

The success of this study would not have been possible without the support and guidance of several individuals. Ms. Kay Albright, Deputy Director of the National Commission for Employment Policy and Project Officer for this study, deserves special credit. Throughout the project, Ms. Albright provided continuing support and guidance. Her commitment to the development of a sound and useful performance management system kept us all focused on maintaining a delicate balance between technical rigor and the practical considerations associated with using performance standards.

Mr. Gregg Knorr of the Employment and Training Administration, U.S. Department of Labor, is also to be thanked for his assistance in obtaining special data runs from the Job Training Longitudinal Survey. Mr. Knorr helped in selecting and defining relevant variables, and in interpreting the tabular results.

At Abt Associates, many individuals contributed to the project. Mr. Glen Schneider assisted with the conduct of our field visits to state and local service delivery areas, and contributed to the development of the site questionnaires. Ms. Jan Nicholson and Ms. Darlene Pinkerton typed and produced the manuscript with their usual good cheer, patience, and high quality work. We thank them for their assistance.

TABLE OF CONTENTS

EXECUTIVE SUMMARY i

1.0 INTRODUCTION TO THE STUDY 1

1.1 Objectives of the Study 1

1.2 Purpose and Organization of the Report 4

2.0 SERVING AFDC RECIPIENTS: ISSUES AND PROBLEMS 6

2.1 Targeting, Program Mix, and JTPA Legislation 7

2.2 The JTPA Service Experience 12

2.3 Key Program Issues 18

2.4 Summary of State Effort to Serve AFDC Recipients 25

3.0 KEY ISSUES IN DEVELOPING A PERFORMANCE STANDARDS SYSTEM 34

3.1 JTPA's Performance Standards System 34

3.2 Developing a Performance Management System 40

4.0 SELECTING VALID PERFORMANCE MEASURES 47

4.1 Performance Measures to be Validated 48

4.2 Estimation of the Validation Benchmarks 49

4.2.1 Impact Estimates From the Original Demonstration Study 49

4.2.2 Construction of Validation Benchmark Measures 51

4.3 Validation Tests Conducted 59

4.4 Key Validation Findings 60

5.0 ESTABLISHING A PERFORMANCE MANAGEMENT SYSTEM 73

5.1 Performance Standards and Creaming: Further Findings from the Validation Tests 75

5.2 Performance Standards, Targeting, and Program Mix: Findings from the Field 84

5.3 Forging a Responsive Performance Management System 89

EXECUTIVE SUMMARY

Performance standards have long been regarded as an important management tool for employment and training programs. They provide basic information for addressing key planning and operational decisions, and allow funding agencies to determine the extent to which programs are moving in a direction consistent with their long-term goals. For the manager who operates within a dynamic, rapidly changing environment, performance standards provide a timely and useful indication of program effectiveness.

In drafting JTPA, both the Congress and the Administration recognized the importance of program outcomes, and established a performance standards system as the cornerstone of the new legislation. In many respects, JTPA was designed as a performance-driven program, intended to clearly link program inputs with program outputs. To achieve this, an elaborate management structure was designed to set numeric performance scores, and provide rewards and sanctions for good and poor performance achievement, respectively.

Although JTPA's performance standards system is expected to fulfill a critical management role, it can also create "perverse incentives" to serve the most job-ready applicants in short-term, low-cost programs. To address these potential effects, JTPA contains at least five mechanisms, including:

1. A requirement that 90 percent of all enrollees be disadvantaged;
2. Targeting provisions requiring specified levels of service to youth, high school dropouts, and welfare recipients;
3. Authorization for Governors to adjust Service Delivery Area (SDA) standards to account for variations in participant characteristics and other factors affecting performance;
4. The establishment of performance standards at a minimally acceptable level; and
5. The availability of financial incentives that Governors can use to influence SDA behavior.

Despite these efforts, there has been growing concern that performance standards have reduced services to the hard-to-employ, and limited the level of investment made in participants. Indeed, this has been argued by practitioners, researchers, and representatives from the Congress. Many believe that the performance standards system has shifted the attention of programs away from long-term employability development to the creation and management of a low-cost, high-volume service delivery system. These concerns have been particularly directed at the Aid to Families with Dependent Children (AFDC) population and especially those identified as hard-to-employ. Despite the pre-

sence of specific targeting provisions, many believe that service providers "cream" from among the eligible population of welfare recipients and provide them with limited, short-term services.

The seriousness of these concerns has caused the Congress and the Department of Labor to reassess the performance standards system critically and consider modifications to it. At the same time, the Congress and Administration are also considering amendments to JTPA that would encourage services to AFDC recipients and other hard-to-employ groups, as well as legislation that would restructure the welfare system and the manner in which training services are provided to public assistance recipients. Since several of the legislative proposals suggest the creation of performance standards to guide the provision of services to welfare recipients, an assessment of such standards is quite important and can provide timely input to current policy deliberations.

This study was designed to assess a series of performance standards issues in the context of employment and training programs targeted on the AFDC population, and especially those who are hard to serve. The study had three major objectives, including:

1. To describe and evaluate the effects of JTPA performance standards on reducing welfare costs, with specific reference to who is served, the services that are delivered, and the effectiveness of programs for welfare recipients.
2. To describe and evaluate the effects of performance standards on service providers in other employment and training programs for individuals similar to those who will be served by the proposed legislative initiatives.
3. To suggest a set of performance measures and a methodology to establish standards for targeted welfare programs, such as those contained in the legislative proposals.

To address these objectives, we conducted two major tasks. The first involved telephone interviews with representatives from JTPA and departments of public welfare in nine states that account for approximately 40 percent of the total U.S. AFDC caseload. These interviews were followed up with field visits to three of the states and the largest urban service delivery area (SDA) within them. In all, the objectives of the interviews were to describe and document: (1) policies and procedures that encourage services to the hard-to-employ; (2) existing welfare performance standards and adjustment procedures; (3) key concerns with performance standards management; and (4) major constraints in serving AFDC recipients, including the hard-to-employ.

The second task included an assessment of the predictive validity of alternative performance measures that could be used for employment and training programs serving AFDC recipients. As part of this task, we also considered methods that could be relied on to set performance standards at the service provider level, including the use of statis-

tical modeling. Five basic groups of performance measures formed the basis of the validation tests.

Candidate Performance Measures Tested

Measure	Time: At Termination	Time: 3 Months Post-Termination	Time: 6 Months Post-Termination
1. Employment Status	x	x	x
2. Average Hourly Wages	x	-	-
3. Average Weekly Earnings	-	x	x
4. Welfare Status	-	x	x
5. Total Welfare Benefit Levels	-	x	x

Ultimately, the quality of the validation tests depends on the quality of the validation benchmarks used. In this study, we relied upon data obtained from the AFDC Homemaker-Home Health Aide Demonstrations. These demonstrations provided training and subsidized employment as homemaker-home health aides to AFDC recipients, who then supplied in-home services to functionally impaired elderly clients. The key advantage of these data is that they were derived from an experimental research design, permitting the estimation of unbiased estimates of program impacts and hence, the validation benchmarks.

Using these data, we constructed two validation benchmarks to assess the validity of the alternative performance measures. These included net gains in earned income and net reductions in public benefits received. Both of these benchmarks are consistent with the explicit goals of JTPA as well as other employment and training programs for the AFDC populations.

Summary of Telephone and Field Findings

Several key findings emerged from our telephone and field interviews with state and SDA representatives. These findings can be summarized as follows:

1. Providing services to the AFDC population effectively was identified as a formidable task by all respondents. While the specific issues raised by our respondents varied, they all agreed that no one agency or program was adequately equipped to meet the employment and training needs of the AFDC population. Four issues, in particular, were consistently cited by both JTPA and welfare department respondents as inhibiting the delivery of effective services including:

   - the difficulty in developing a coordinated strategy between JTPA, welfare, social services, and education at both the state and local levels;

- the limited and restrictive nature of support services and other benefits, such as health insurance and housing;
- the work disincentives imposed by current AFDC earnings disregard provisions and the quality of available employment opportunities; and
- political pressure on programs to show quick, visible evidence of success (e.g. high placement levels at relatively low costs).

2. In partial response to these issues, each of the states we visited had initiated their own brand of welfare reform designed to strengthen the relationship between work and welfare. While these initiatives varied in their stage of development and specific program approach, they shared the common feature of attempting to broaden and enrich the typical mix of employment and training services in an effort to improve their responsiveness to individual needs.

3. While all states indicated that they had taken steps to target on the hard-to-employ AFDC recipient, there was substantial diversity in how such individuals were defined. In some cases, the hard-to-employ were defined in terms of teen parents, while in others, they were identified on the basis of prior levels of welfare dependence. What was most clear, however, was the absence of unanimity regarding the characteristics of the hard-to-employ. Sharp variations were found at the state as well as at the local level.

4. Performance standards were also cited as inhibiting the delivery of services to AFDC recipients, although they were not among the most important factors cited by respondents. Two issues, in particular, stood out during our interviews.

   The first focused upon the cost standards used in JTPA and in many of the employment and training programs sponsored by state departments of public welfare. At both the state and local SDA levels, the pressures to keep costs down were viewed as limiting the types of program investments many believed were necessary to help AFDC recipients achieve stable and rewarding employment. In fact, in several of the state JTPA offices that we interviewed, efforts were underway to either eliminate the cost standards or substantially raise them above established levels.

   Perhaps more important was the limiting influence many believed that performance-based contracts imposed on the enrollment and investment decisions made by program vendors. These contracts establish specific unit costs for individual performance benchmarks, such as participation rates and program outcomes. For the SDAs we visited, performance-based contracts were viewed as an efficient tool for ensuring accountability and a focus on tangible, positive outcomes. For contractors, however, they were viewed as representing an important source of conflict when viewed in the context of matching services to the needs of AFDC recipients. Since contractor reimbursement was tied to the achievement of established performance benchmarks, and because available adjustment tools were not typically administered at that level, vendors were often caught between the goal of providing effective services and the necessity of maximizing the likelihood of cost reimbursement. In many cases, this was believed to provide incentives to cream from among the eligible population.

Beyond these results, we also conducted a careful review of the JTPA legislation and available evidence on the adequacy of service delivery to AFDC recipients. Our findings pointed up two important results. First, the most recent evidence on service shares to the AFDC population indicate that JTPA has indeed met its targeting requirement of serving AFDC recipients in proportion to their share of the eligible population. This suggests that current concerns over creaming from among AFDC recipients must relate more to the characteristics of recipients served and not the relative service share received by them. However, when placed within the context of the JTPA legislation, it is not clear that certain types of individuals are being excluded from receiving services that they otherwise should have received.

Several features of the JTPA legislation were identified as potentially contributing to the concern over the types of individuals being served. These include:

1. The strong emphasis of JTPA on local determination of needs and priorities.
2. The absence of explicit targeting provisions regarding the types of AFDC recipients to be served.
3. The limited availability of resources for support services and remedial education.
4. The responsibility to coordinate with other related programs, but the lack of authority to do so.

Taken together, our findings indicated that changes to the existing performance standards may be a necessary but not a sufficient condition for stimulating the delivery of more effective services to the AFDC population, and particularly the hard-to-employ. Other more basic issues, ranging from institutional impediments to legislative ambiguity, may also have to be addressed if employment and training programs are to become more targeted on and responsive to the needs of the AFDC population.

Summary of Key Validation Results

A key result that emerged from our validation analysis is that there are candidate measures that represent valid predictors of both earnings gains and reductions in welfare dependency. Of over twenty separate correlations evaluated, fourteen or 70 percent were found to be significant and with the expected sign. As was anticipated, however, some measures performed better than others.

In particular, we found that post-program measures performed substantially better than those taken at the time of termination; and of the post-program indicators, those measured over six months had greater predictive validity than those calculated three months after program termination. We also found that earnings gains were considerably easier to predict than reductions in welfare dependency. While all but two

of the candidate indicators were valid predictors of earnings gains, five of the candidate measures were uncorrelated with the welfare reductions benchmark. For the most part, the employment and earnings indicators did not perform as well with net reductions in welfare benefit levels as they did with earnings gains.

There were, however, two measures that appeared to be valid predictors of both benchmarks. Average weekly earnings and average monthly benefit levels measured over the six months following termination were found to be the best predictors of both benchmarks. These measures were also found to be well correlated with each other, suggesting that they are complementary rather than competitive.

This is not to say that the three months measures are invalid. To the contrary, employment status, average weekly earnings, and welfare status and benefit levels at three months were each significantly correlated with the earnings benchmark. While the magnitude of the correlations is smaller than that at six months, the difference tended to be relatively small. The same is true about the correlations between the two welfare indicators and the welfare reductions benchmark; while six months out performs the three month measures, the difference are not always large.

The only exception is the correlation between the employment and earnings indicators and the welfare benchmark. Here, the three month measures are not only smaller than the six month indicators, but they are also statistically insignificant.

Because there is often interest in using more than one performance measure, we also assessed the incremental validity of using one indicator in conjunction with another. This test was performed using one's placement status in conjunction with other indicators. Placement status was selected as the base indicator because of the current and historic preference given to it in employment and training programs. Overall, our findings indicated little support for using placement status in conjunction with other indicators. For one thing, placement status, by itself, was not found to be a valid predictor of net reductions in welfare dependency. In addition, while placement status did predict earnings gains, its predictive power was relatively low and added little information not already provided by the post-program measures.

Given these results, what can we say about the choice of performance measures for employment and training programs targeted on AFDC recipients? From the point of view of statistical validity alone, the choice is clear. Average weekly earnings and total welfare benefit levels over the six months following program termination consistently performed better than those indicators measured at three months. Moreover, both of these measures were highly correlated with the two validation benchmarks, indicating their relevance for both earnings gains and reductions in welfare dependency.

However, other more practical considerations must also be taken into account when selecting appropriate performance measures. For one thing, the ease and cost of data collection is also important. To the extent that a candidate measure requires relatively complex computations or the collection of difficult-to-obtain data, its usefulness will be limited by the added difficulty it presents to managers and administrators. Similarly, if the post-program period is too long, then the value of increased predictive validity will be offset by the limited utility of the indicator on an ongoing basis.

These considerations may serve to question the "statistical" appeal of earnings and welfare benefit levels measured over the six months following program termination. While both indicators have the greatest predictive validity, they are relatively difficult to measure, and require comparatively long lead times to use for monitoring program performance. For this reason, it is appropriate to also consider their three-month counterparts as well as welfare status and/or employment status at three months. These measures did not perform as well as the six-month measures, but the observed differences were small. From a practical point of view, the tradeoff between predictive validity and practical utilization would appear to be one worth considering.

Ultimately, the choice of final performance measures will have to balance these practical considerations with the predictive validity of the alternative indicators. While difficult tradeoffs may have to be made, the findings in this report provide ample choices of valid candidate measures from which to choose.

Forging an Effective Performance Management System

The development of a flexible performance management system needs to go beyond the identification of statistically valid indicators. At least two additional issues warrant attention. As mentioned above, the first is that of feasibility. The indicators selected for use must be retrievable in a timely and accurate fashion and in a manner that does not place substantial burden or cost on those responsible for administering them.

In this study, the merits of six month indicators must thus be weighed against these considerations. Although longer post-program periods produce more valid performance measures, they limit the use of the indicators on a timely basis and may also contribute to greater inaccuracies if participant surveys are relied on to obtain the data. Ultimately, these choices can be based only partly on empirical evidence. Policy makers and program administrators will have to rely on their own judgments regarding costs and timeliness.

Second, because the existing performance measures have been viewed as providing creaming incentives, it is also important to consider this possibility and how it might be curtailed in the establishment and management of performance standards. Moreover, even if creaming per se does not appear to be a problem, it still remains to devise a methodology for setting performance standards in an objective and equitable manner.

Our first step in addressing this issue was to examine the relationship between each of the candidate performance measures and selected background characteristics that could be used for selective enrollment, such as prior work and welfare dependency history. This analysis also provided the benefit of identifying key client variables that would be important to consider in statistical modeling.

Overall, our findings did indicate that the traditional employment, earnings, and welfare status performance measures do have the potential for introducing creaming incentives. For the most part, the candidate measures included in this study were highly correlated with individual background characteristics. Since program operators can readily select individuals on the basis of these types of personal indicators, it follows that a concern over performance achievement, especially in light of performance-based contracts, can lead to enrolling those individuals with the greatest likelihood of achieving established benchmarks.

One way to address this is through statistical modeling. Indeed, the intuition underlying the Department of Labor's adjustment model is that by removing the effects of client characteristics on performance achievement, SDAs (and program operators) will have limited incentives to engage in selective enrollment.

Whether or not this approach is effective has come under increasing scrutiny. The states and SDAs we talked with indicated that the available adjustment models have only a limited influence on targeting and program mix decisions. The small size of the adjustment factors and the relative ease by which SDAs can meet their performance standards were cited as contributing to this perception. However, beyond reasons related to the technical merits of the model, an additional consideration is that using the model, in a practical planning sense, is strictly voluntary. In other words, an SDA must decide first to target on AFDC recipients before it can take advantage of the adjustments provided by the model.

Compounding this is the problem faced by service vendors. It is at this level that actual enrollment decisions are made, and where creaming pressures were most evident. Since the adjustments currently offered to SDAs are not directly passed onto vendors, contractors operate largely with unadjusted performance scores. To the extent that they are motivated to achieve high levels of performance through, for example,

performance-based contracts, they have an incentive to cream from among the eligible population.

Thus, the availability of statistical adjustment models alone will not necessarily stimulate service delivery to the AFDC population or subgroups of it. Moreover, while it is possible to improve the modeling process through the inclusion of additional variables, such efforts should not be expected to produce large changes in current behavior.

There are, however, certain steps that can be pursued in conjunction with statistical modeling. One would be to increase the incentives service providers have in serving AFDC recipients. Most of the state and SDA respondents we spoke with viewed JTPA's six percent incentive funds as a viable mechanism to induce certain types of targeting and program mix decisions. However, they also cited the need to better direct the use of these funds to targeting issues and to also increase the funding level allocated to incentive dollars.

The second would be to make the statistical adjustment models available to local vendors. Currently, this is not the case. While SDAs rely on the models, they do not and cannot make them readily available to their vendors during the planning process. As discussed earlier, this leaves program operators without the benefit of the adjustment models, limited as they may be.

Vendor use of the adjustment models could be accommodated in one of two ways. On the one hand, SDAs could include the adjustment factors from the existing models in vendor RFPs as an optional (or required) planning and bidding tool. On the other hand, locally-based models, using SDAs micro-data could be developed and used in the same fashion.

A third a final step would be to increase the level of technical assistance provided to states and particularly SDAs regarding the provision of effective services to the AFDC population. Our results indicated, for example, that assistance in the area of identifying the hard-to-employ population would help JTPA to better target its resources and distinguish among the various types of (AFDC) participants it potentially could serve. Similarly, technical assistance in developing appropriate program designs and coordinating with other agencies to leverage resources would also be quite helpful, given the difficulties we observed in the area of interagency coordination.

Taken together, these types of changes to the current performance standards structure should be expected to provide some relief from existing creaming incentives. However, as discussed earlier, absent selected changes in the targeting and program provisions of JTPA, they will not necessarily produce large changes in state and SDA choices over who to serve and in what programs.

If there is an interest in using the performance standards system to stimulate greater service levels to AFDC recipients, it would appear that some other, or additional approach may be warranted. On the basis of our validation research, one viable option would be to set target group performance standards. Short of amending the targeting provisions on the JTPA legislation, target group performance measures would exert a clear pressure of the state and local levels to provide services to selected population subgroups.

There are two ways in which target group performance measures could be established. One would be to set service shares for pre-specified groups. Since JTPA already contains service share requirements for the AFDC population, the use of such performance measures would have to focus on subgroups of the welfare population, such as long-term recipients, teen parents, or those with very limited work histories. In view of the wide range of definitions currently in use by the states to identify hard-to-employ AFDC recipients, such efforts would likely have to consider a range of options and result in the selection of one that best meets the interests of the states and their SDAs.

Another option for setting target group performance measures would be to establish specific outcomes for either the AFDC population in total, or for selected subgroups of it. In this case, the focus would be placed on achieving pre-established levels of performance, in contrast to service shares. The advantage of this option is that it maintains JTPA's focus on outputs, as distinct from process.

But the use of service share measures would not have to replace the current emphasis on program outputs. Rather, they could be viewed as a first-order condition of performance. In other words, employment and training programs would be required first to achieve certain levels of service to pre-specified target groups. Having achieved this, they would then be judged on the basis of how well they performed in terms of program output measures. Thus, modeling would continue, but it would not be viewed as the only vehicle for encouraging services to particular groups that would otherwise be avoided due to concerns over performance achievement.

Adopting such an approach also raises certain issues. For one thing, creating target group performance measures could be viewed as running counter the basic principles of decentralization underlying the JTPA program. Second, the performance management system would have to be provided with the support necessary to help overcome existing barriers to coordination, and identify as well as implement the most effective way to organize and provide services to the special population groups. Absent these efforts or new legislation to support them, our findings indicated that even the best performance management system will not alter the behavioral preferences of program operators.

1.0 INTRODUCTION TO THE STUDY

1.1 Objectives of the Study

This is the final report of a study on the role and use of performance standards for employment and training programs serving welfare recipients.[1] The study had three major research objectives, including:

1. To describe and evaluate the effects of JTPA performance standards on reducing welfare costs, with specific reference to who is served, the services that are delivered, and the effectiveness of programs for welfare recipients.

2. To describe and evaluate the effects of performance standards on service providers in other employment and training programs for individuals similar to those who will be served by the proposed legislative initiatives.

3. To suggest a set of performance measures and a methodology to establish standards for targeted welfare programs, such as those contained in the legislative proposals.

Our approach to these objectives consisted of three tasks. The first task was to assess the influence of performance standards on targeting, program mix, and investment decisions. Here, we conducted a qualitative analysis of the role of performance standards in planning and providing services to AFDC recipients, and particularly those that are hard-to-employ.

The analysis was based upon telephone interviews with the nine states that account for the largest share of AFDC recipients in the U.S., and field visits to a sub-set of three states. In each state, interviews were conducted with representatives from both JTPA and welfare departments, while during the field visits we also visited the largest urban SDA. The states and SDAs included in the study were:

State Included in Study	Telephone Interview	Field Visit	SDA
Florida	x	x	Miami
Illinois	x		
Michigan	x		
New Jersey	x		
New York	x	x	New York City

[1] An interim study report was published in May, 1988. See: Zornitsky, Jeffrey, Rubin, Mary, Serving AFDC Recipients: Initial Findings on the Role of Performance Standards, National Commission for Employment Policy, Washington, D.C., May, 1988.

State Included in Study	Telephone Interview	Field Visit	SDA
Massachusetts	x	x	Boston
Ohio	x		
Pennsylvania	x		
Texas	x		

The telephone survey was designed to provide descriptive information regarding welfare performance standards and provision of employment and training services to hard-to-employ individuals. In each state, two interviews were conducted: one with the manager of JTPA performance standards, and another with the individual responsible for employment and training in the state welfare department. These two respondents were asked a series of brief questions designed to describe and document: 1) policies and procedures that encourage services to hard-to-employ individuals; (2) existing welfare performance standards and adjustment procedures; (3) key concerns with performance standards management, with special reference to the mix, level, and setting of standards; and (4) major constraints in serving the hard-to-employ. To the extent available, respondents were also asked to mail relevant documents, policies, and performance related data to support our effort.

The information obtained from our telephone survey was used for two purposes. First, we summarized key findings for each individual state, highlighting differences between JTPA and non-JTPA programs. Second, we conducted a cross-site analysis on a topic-by-topic basis in order to portray how states have addressed key issues related to welfare performance standards and the delivery of services to the hard-to-employ. Results from interviews in the field are designed to enrich these discussions.

The SDA site visits were intended to illuminate the interaction of performance standards with key service delivery policies and procedures. The primary areas of inquiry guiding the field visits included: (1) policies and practices used to shape recruitment and selection, (2) criteria for selecting a program mix and matching individual needs with services, and (3) constraints imposed by performance standards.[1]

The second project task was to conduct a series of validation tests to assess the predictive validity of alternative performance measures using experimentally derived

---

[1]For a detailed discussion of the methodology used see: Rubin, Mary and Zornitsky, Jeffrey, Serving AFDC Recipients: Initial Findings on the Role of Performance Standards, National Commission for Employment Policy, Washington, D.C., May 1988.

impact estimates obtained from the AFDC Homemaker-Home Health Aide Demonstrations. These demonstrations provided training and subsidized employment as homemaker-home health aides to AFDC recipients, who then supplied in-home services to functionally impaired clients. Thus, the demonstrations sought to alleviate a major social problem, that being welfare dependency. The Health Care Financing Administration (HCFA) was the lead agency for implementing the demonstrations in seven states: Arkansas, Kentucky, New Jersey, New York, Ohio, South Carolina, and Texas. Operations began in January 1983 and ended June 30, 1986. Participation in both the training and service components of the demonstrations was voluntary. To permit rigorous evaluation of demonstration effects, eligible applicants (both potential trainees and potential clients) were randomly assigned in equal numbers to treatment and control groups.

At the time of enrollment, both groups were administered a baseline questionnaire to identify their prior welfare and employment histories as well as their personal characteristics. Following baseline, treatment and control group members were subject to one followup interview, administered between 5 and 32 months later; on average, the followup period was 12 months long. The followup data, combined with welfare payment information obtained from state administrative records, provided the basis for estimating (experimental) impacts on welfare dependency, employment, and earnings. Thus, these data offer a unique opportunity to validate welfare performance measures with experimentally-based impact estimates.

In selecting these data to conduct the validation tests, we realized that the demonstrations represent one specific approach to training and placing welfare recipients. While it might have been preferable to perform the tests on a broad range of employment and training programs for welfare recipients, such data are not available. Thus, we viewed the AFDC Homemaker-Home Health Aide Demonstrations as the best choice for four key reasons, including:

1. The data focus on the specific population of interest, namely AFDC recipients. On average, trainees served in the program had not worked for 34 months prior to entry, were 31 years of age, cared for two dependent children, were black or Hispanic (72 percent), and had not gone beyond high school (78 percent). In many respects, they resemble the target group of interest to this study.

2. The data permit estimation of experimentally-based welfare reduction impacts.

3. The data contain detailed information on prior work history and welfare status.

4. The findings can be combined with those obtained previously and with the research results produced by MDRC to yield a comparative framework for assessing the validity of alternative performance measures.

Using these data, we constructed two validation benchmarks -- impacts on welfare reductions and earnings -- and tested the validity of a number of alternative measures that vary by both definition and measurement. The measures we tested focused on: 1) employment status; 2) earnings; 3) return to school/enrollment in training status; and 4) welfare status. Each measure was constructed at the time of program termination as well as over a three- and six-month post-program period.

To judge the relative merits of alternative measures, we relied on three criteria. The first was face validity; that is, the measures will have to be related to the purpose of the program and be viewed as having intuitive appeal. The second criteria was feasibility. Here, our concern was with the practical difficulties associated with obtaining state and local data. The performance measures selected must not present service providers with undue costs or problems in collecting and reporting the necessary data. To address the last criterion -- predictive validity -- we relied on three tests.

First, we examined simple correlations between each candidate measure and the validation benchmarks. The relative magnitude of the resulting correlations thus provide an indication of the relative validity of each measure. In practice, however, more than one measure may be used. As a result, we also conducted joint and incremental tests of validity. Alternate groups of indicators are assessed against each other, in an effort to determine by how much validity increases when selected measures are added to another.

The success of a performance management system is very dependent on the framework adopted to adjust standards for differences in client characteristics and local economic conditions. Thus, our third task examined the types of adjustment models that would be most appropriate for targeted welfare training programs. Focusing on the most promising candidate performance measures, we relied upon multivariate statistical techniques to determine those potential adjustment factors that best explain variations in each indicator.

1.2 Purpose and Organization of the Report

This final report is designed to achieve two key objectives. First, it relies on the results from our telephone interviews and field visits to summarize key issues and problems in serving welfare recipients, including the role of performance standards.

Second, the report presents the results obtained from our validation research and their implications for the design of a performance management system.

Chapter 2 of the report summarizes key issues and problems involved in the provision of employment and training services to welfare recipients. Chapter 3 discusses the development of a performance standards system. Chapter 4 discusses the selection and validation of performance benchmarks, while Chapter 5 draws on key findings from the validation tests and from the field interviews in order to suggest how the performance management system could be made more responsive to the key issues and problems discussed in earlier chapters.

2.0 SERVING AFDC RECIPIENTS: ISSUES AND PROBLEMS

In the 1980s welfare reform has again become an issue of national prominence. Ongoing debates at the federal and state levels have resulted in a broad consensus across the political spectrum that work opportunities must be an integral part of the welfare system, and that the system of public welfare in the United States must be structured to promote employment rather than dependency. On the basis of our telephone interviews and field visits, it is evident that there is substantial activity aimed at pursuing this objective. Most of the states and SDAs we interviewed are engaged in implementing ambitious efforts to target services to those most in need (individuals with multiple barriers to employment) and to deliver a broad array of basic educational, vocational skill training, job placement, and support services. Yet within this, there remains concern that performance standards constrain the choices service providers face in determining the level and mix or programs to offer AFDC recipients.

These concerns are being increasingly voiced, at least in the regular, ongoing employment and training programs, leading many policy makers to conclude that the performance standards need to be more carefully managed and modified to better stimulate service delivery to AFDC recipients, and especially the hard-to-employ. This perspective is particularly evident in the JTPA program, given its emphasis on performance management and its central role in providing employment and training services to the disadvantaged population. The Senate Committee on Labor and Human Resources probably best captured this sentiment in its report accompanying the proposed Jobs for Employable Dependent Individuals Act:

> The Committee conducted hearings on January 21, February 3, February 4, and March 6, 1987 to receive testimony on both the effectiveness of the Job Training Partnership Act in serving long-term welfare recipients, and the specific provisions and ramifications of S.514. Witnesses commented on the extent and nature of long-term welfare dependency, particularly with regard to recipients of Aid to Families with Dependent Children (AFDC), and on those aspects of JTPA's program design which contribute to "creaming." This practice has serious consequences for long-term welfare recipients who seek employment. Creaming is the practice of preferential enrollment of those among the eligible population who require the least assistance to prepare for a job. The present JTPA performance standards encourage creaming by rewarding the SDAs with the lowest training and placement costs per individual. As a result, individuals needing more extensive or costly training are neglected or underserved.[1]

---

[1] Committee on Labor and Human Resources, U.S. Senate, Jobs for Employable Dependent Individuals Act: Report Together with Additional Views, Report 100-20, 100th Congress, Washington, DC, March 20, 1987.

The implication of this is that if performance standards are modified, or in an extreme case eliminated, service delivery to long-term AFDC recipients would be substantially improved. Disincentives to serve this subgroup of the population would be eliminated, targeting would be redirected, and program services would be modified to accommodate the needs of a relatively different client population. The JTPA program essentially would shift its focus voluntarily to a more narrow subset of the eligible population and make those program design changes necessary to service it effectively.

Unfortunately, there is little evidence that this outcome would occur. While studies of the JTPA program have shown some evidence that service providers engage in activities that promote selective enrollment, it is not clear that they are motivated chiefly by the desire to achieve performance targets. There are other factors that may also be shaping the types of targeting and program investment decisions made by service providers.

This chapter of the report is designed to create the program and policy context for understanding how performance standards may affect the delivery of services to the welfare population. Since performance standards are managed within a complex and evolving set of programs, it is important to identify the key factors that influence whom they serve and how they are designed. In this way, it is possible to better develop a performance management system that supports the delivery of services to AFDC recipients.

We begin the chapter with a discussion of the structure of JTPA and its consistency with the objective of serving AFDC recipients, particularly those that are hard-to-employ. We then turn to a presentation of the recent JTPA service experience, relying on data derived from the Department of Labor's Job Training Longitudinal Survey. In the final section, we review the key program and policy issues that emerged from both our telephone interviews and field visits.

2.1 Targeting, Program Mix, and JTPA Legislation

In 1986, $17,757 million of public funds were distributed in the form of cash benefits to AFDC recipients.[1] While there has been a long-standing general agreement that such public monies should be used to meet the subsistence needs of these recipients

[1]Statistical Abstract of the United States, 1987.

and their families, there are now growing concerns that these public resources should also be directly linked to efforts that enable a family to establish financial stability. Such concerns are being articulated at all levels of government and have been manifested through the introduction of several federal welfare reform proposals. Common to each of these proposals are strategies that would enable individuals on public assistance to obtain an interim mix of work and welfare while working towards the ultimate goal of leaving the public assistance rolls altogether.

The Job Training Partnership Act was established in 1983 to meet the employment and training needs of economically disadvantaged individuals, including AFDC recipients. In providing these services, the Act was designed to achieve two goals: (i) to increase employment and earnings; and (ii) to reduce welfare dependency. While the goal of reducing welfare dependency established at least an indirect targeting link between JTPA and AFDC, the Act goes on to target welfare recipients specifically; that is, Section 203(b)(3) of the Act requires localities to serve these individuals in a manner which is proportional to their representation within the eligible JTPA population. In this way, then, the JTPA system was designed to meet the employment and training needs of AFDC recipients within the broader confines of the economically disadvantaged population. It is, however, important to point out that the JTPA funding base to accomplish these objectives is substantially smaller than that of AFDC. In FY1986, $1.782 million was allocated to JTPA services; this amount is about one-tenth of total cash assistance distributed through the AFDC program.[1]

Available evidence indicates that JTPA has indeed responded to this targeting mission. A recent study of JTPA service patterns has revealed, for example, that AFDC recipients have been served by JTPA in a manner proportionate to their representation within the JTPA-eligible population.[2] In comparison to an 18.1 percent share of the eligible population, AFDC recipients accounted for 21 percent of all JTPA participants during Program Years 1984 and 1985. Yet, despite this, the Congress and other policymakers have raised concerns that JTPA is not responding to the employment and training needs of the hard-to-employ subgroups of the eligible population, particularly welfare recipients.

---

[1] Catalog of Federal Domestic Assistance, FY1986.

[2] Sandell, S.H. and Rupp, K., Who served in JTPA Programs: Patterns of Participation and Inter-group Equity, National Commission for Employment Policy, Washington, DC, February, 1988.

It would seem that if JTPA is truly not fulfilling its legislative mandate to serve the AFDC population, then the concerns over "creaming" and other targeting inequities must relate to the characteristics of the welfare recipients who are receiving JTPA services, and not to the relative service share received by the AFDC population. That is, while JTPA may be serving an appropriate share of AFDC recipients, it may not be targeted on those who are considered to be hard-to-employ, because of their limited skills, weak labor force attachment, and/or long-term welfare dependency. However, criticisms of JTPA for a failure to provide services to such individuals may not be fully appropriate. Several factors shape the client and program mix at the local level.

For one thing, a distinguishing hallmark of JTPA is the strong emphasis on localized control over service delivery decisions. By design, the JTPA program encourages states and localities to make program and targeting decisions that best reflect their own perceptions of needs and priorities. These priorities may not always be fully consistent with existing or new national priorities, but they are consistent with the structure of the program and its emphasis on local discretion. The outcomes of state and local decision-making may thus be viewed as the aggregate effect of a program that encourages diversity in state and local choices over whom to serve and what programs to offer. It may be the case, however, that such a decentralized structure fosters a "mismatch" between the targeting and program priorities perceived at the national level and those developed at the state and local levels.

Second, the JTPA program is not explicitly targeted on individuals in hard-to-employ subgroups. Nowhere in the Act is this group explicitly mentioned or precisely defined. Since JTPA was deliberately designed to provide targeting flexibility, SDAs can fulfill their responsibilities under the Act in a manner consistent with their own perceptions of need. The lack of specific within-group targeting as well as JTPA's intentional systemic flexibility both contribute to the growing mismatch between public expectations and JTPA's legislature structure. Absent more specific targeting requirements, it is difficult to argue that some groups, such as the hard-to-employ, are being excluded from receiving JTPA services that they otherwise should have received.

As an example, consider the requirement in Section 141(a) that "each job training plan shall provide employment and training opportunities to those who can benefit from and who are most in need of such opportunities." Because the interpretation of this requirement is left to local discretion, its application can be expected to result in a wide range of target group and program mix selections. Moreover, it can be shown that certain interpretations of those who can benefit from and are most in need of services are

inconsistent with popular perceptions of the characteristics of the hard-to-employ.[1]

For instance, one way to characterize the hard-to-employ is in terms of labor force status. Since the problem of disadvantagedness tends to be related more to a lack of labor force participation than low earnings or high unemployment, it is reasonable to say that the hard-to-employ are concentrated among those who report being out of the labor force.[2] Similarly, since the majority of public assistance resources are consumed by a relatively small proportion of recipients who experience long, continuous spells and have weak attachments to the labor market, it is also reasonable to assume that among the AFDC population, the hard-to-employ are, more often than not, out of the labor force.[3] While these groups are certainly among those who are most in need of services, their weak attachment to the labor force can be used as the basis to argue that they cannot readily benefit from program participation.

Since program enrollment represents one vehicle to secure employment, weak labor force participation suggests little interest in seeking out services and hence, a limited ability to benefit from employment and training programs. It follows then, that the unemployed might best fit the meaning of "benefit from and most in need of" services because they are among the eligible population (or subgroups thereof) and, by virtue of their job search behavior, have a relatively high level of interest in program participation. To the extent that service providers adopt this perspective, their actions can be interpreted as logical and consistent with the discretion offered by JTPA, though not necessarily with the objectives of serving the hard-to-employ.

Other factors may also shape the targeting and program mix decisions made by service providers. These include institutional impediments to coordination between JTPA and welfare departments, limitations on the use of funds for support services and

---

[1]See Sandell and Rupp, op. cit. for a fuller treatment of this argument.

[2]For a description of the labor force status of the disadvantaged see: Glen Schneider et al., An Assessment of Funding Allocation Under the Job Training Partnership Act, Abt Associates Inc., Cambridge, MA, 1986.

[3]See, for example: Mary Jo Bane and David T. Ellwood, Slipping Into and Out of Poverty: The Dynamics of Spells, Harvard University, Cambridge, MA, August 1985; Mary Joe Bane and David T. Ellwood, The Dynamics of Dependence: The Routes of Self-Sufficiency, Urban Systems Research and Engineering, Inc., Cambridge, MA, June 1983; and David T. Ellwood, Targeting "Would-Be" Long-Term Recipients of AFDC, Mathematica Policy Research, Inc., Princeton, NJ, January 1986.

stipends, decreasing levels of the funding, and limited resources for remedial education. Taken together, these factors may well inhibit the ability of service providers to target resources on AFDC recipients, and offer them comprehensive, responsive services. Moreover, a second set of decisions which influence who JTPA is intended to serve and who is actually served involves participation choices made by welfare recipients themselves. For many AFDC recipients, these choices are strongly influenced by the availability of supportive services because of the need to juggle training and family responsibilities as well as jobs paying reasonable wages.

Within this context, one JTPA official we interviewed explained his view of the problems associated with serving AFDC recipients:

> "A public agency like an SDA is really between a rock and a hard place. It feels all sorts of pressure to help fill local service and retail job openings; however, such jobs pay poorly and won't be sufficient to get people off of welfare. Therefore, SDA staff think that they need to answer to the employer community, but also know that they may not be doing any good to help welfare clients.
>
> Furthermore, you can't expect welfare recipients to just give up public benefits, no matter how bad they want to work. It is essential that the jobs they are placed into are worthwhile and pay decent wages. This is a situation in which the government must act by providing supplemental benefits so that this conflict will not keep welfare recipients from leaving the rolls and getting jobs."

Within the salience of these institutional factors, there is also the possibility that JTPA's emphasis on performance achievement constrains the choices service providers face in determining who to serve and the types and levels of services to offer. To the extent that SDAs and their vendors are concerned with performance achievement, it is quite feasible that they will choose a participant and program mix most conducive to meeting their expected performance outcomes. These choices may well be inconsistent with serving the hard-to-employ since by definition, these individuals are among those least likely to register high scores on JTPA's existing performance standards.

Taken together, these attributes of the JTPA program suggest that it may not be well positioned to dedicate its targeting and services to the hard-to-employ. Performance standards notwithstanding, the decentralized nature of the program coupled with the targeting flexibility it grants to states and SDAs suggests that legislative changes may be needed to effectively redirect the activities of service providers.

Based on our interviews with staff from both JTPA and welfare, it appears that there is little consensus on how to define the hard-to-employ and the types of services

that would best meet their employment and training needs. As we discuss later in this chapter, state-initiated welfare reform efforts display a noticeable degree of variation in the criteria used to define the hard-to-employ AFDC recipient. Moreover, there is widespread agreement that a distinguishing characteristic of hard-to-employ individuals is the presence of severe barriers to employment, including but not limited to illiteracy, teen parenthood, weak attachment to the labor force, and a pattern of long-term dependency on public welfare. As would be expected, this has resulted in a wide range of program models that states have adopted as part of their effort to better link work and welfare.

2.2 The JTPA Service Experience

Policy makers and practitioners concerned about work and welfare frequently suggest that the mix and intensity of JTPA services may be inadequate to address the employability needs of the AFDC population, and particularly those that are hard-to-employ. Specifically, AFDC mothers are viewed as requiring long-term services that are both intensive and comprehensive so that they can maintain the motivation necessary to address their employment barriers and eventually, achieve successful labor market outcomes.

While is is difficult to assess the responsiveness of JTPA to the needs of the AFDC population, it is possible to examine the pattern of services provided by JTPA and how they vary between AFDC recipients and those not receiving public assistance. In this section of the chapter, we present selected information on the types of services provided to AFDC recipients, and the outcomes they experience as a result of program participation. The data have been derived from the Job Training Longitudinal Survey for program years 1984 through 1986.

Exhibit 2.1 and 2.2 present information on the types and intensity of services provided to AFDC recipients and all individuals served by the program. Several interesting findings emerge. First, it is evident that the JTPA program provides relatively short-term services to all of its participants, regardless of the type of service received. On average, individuals who terminated from JTPA between program year 1984 and 1986 received services for 100 days, or approximately 3.3 months. While there is variation around this average with respect to the type of service provided, the median duration never exceeds 4.4 months.

While AFDC recipients do tend to receive services for longer periods of time than the average terminee, the additional time spent in the program represents less than

EXHIBIT 2.1

MEDIAN LENGTH OF STAY OF JTPA TERMINEES,

BY TYPE OF AFDC RECIPIENT, PY84-PY86 (in days)

		AFDC Recipients		
	All Terminees	Total	Parents	Female Single Parents
	1,937,844 (100.0)	407,956 (21.1)	274,610 (14.2)	220,109 (11.4)
Total Terminees	99.7	114.5	116.0	117.3
Type of Training				
Job Search Assistance	27.1	49.5	46.2	45.2
CT/Occ. Skills	130.8	141.0	144.0	143.1
OJT	100.8	102.5	100.4	95.5
Work Experience	128.5	140.0	114.2	120.7
Other	108.8	119.2	128.9	129.9

Source: Job Training Longitudinal Survey, Unpublished Data

EXHIBIT 2.2

TYPES OF SERVICES RECEIVED BY JTPA TERMINEES,

BY TYPE OF AFDC RECIPIENT, PY84-PY86

	All Terminees	AFDC Recipients		
		Total	Parents	Female Single Parents
	1,937,844 (100.0)	407,956 (21.1)	274,610 (14.2)	220,109 (11.4)
Type of Training				
Job Search Asst. excl.	19.0	15.7	16.4	16.0
CT/Occup. Skills	23.1	31.0	36.7	38.9
OJT	23.2	14.1	15.7	13.6
Work Experience	8.5	8.6	3.7	3.7
Other	26.1	30.6	27.5	27.9
Received Basic Education				
Yes	10.0	14.5	14.2	15.1
No	90.0	85.5	85.8	84.9
Type of Basic Education				
Basic Education Only	5.1	7.8	7.3	8.0
Basic Ed./CT-Occ. Sk.	1.4	1.8	2.2	2.2
Basic Ed./OJT	0.4	0.5	0.5	0.5
Basic Ed./Other	3.2	4.5	4.3	4.5
Received Support Serv.				
Yes	16.6	20.8	22.4	22.8
No	83.4	79.2	77.6	77.2
Type of Support Serv.				
Child Care	0.8	2.4	3.4	3.9
Transportation	7.0	9.4	11.3	11.4
Needs-Based Pay.	4.7	5.4	5.1	5.0
Other	9.4	12.0	13.2	13.7

Source: Job Training Longitudinal Survey, Unpublished Data

one month. As Exhibit 2.1 shows, the median duration of participation for all AFDC recipients was 115 days or 3.8 months.

Despite this, AFDC recipients do tend to receive a somewhat different mix of services than other participants (Exhibit 2.2). As a group, those collecting AFDC payments are more often served in occupational training and other specialized programs than their counterparts. Moreover, they receive basic education more frequently and, as would be expected, are more likely to receive supportive services, regardless of the type.

To some extent, these differences reflect the fact that AFDC recipients served by JTPA have greater educational deficits than other participants. As Exhibit 2.3 shows, 31 percent of all AFDC recipients were classified as high school drop-outs, compared to 27 percent of all terminees. Interestingly, however, roughly two-thirds of both groups were reported as unemployed at the time of program application. If, as has been shown elsewhere, AFDC recipients are more often out of the labor force than either unemployed or employed, then these data would suggest some form of selective enrollment in the JTPA program.[1] In viewing this, however, it is important to be aware of likely inconsistencies between the formal definition of unemployment used by the Bureau of Labor Statistics (BLS) and that employed by intake staff at the local SDA level. In many cases, SDA staff simply ask prospective enrollees whether they have been looking for work or if they are interested in working; an affirmative response is often recorded as being unemployed. This stands in contrast to the rigorous criteria used by the BLS in determining one's labor force status.

Even if SDAs selectively enroll AFDC recipients, they are not able to serve this group as effectively as others. As Exhibit 2.4 shows, AFDC recipients experience lower placement rates and hourly placement wages than their counterparts. In comparison to a 54 percent placement rate for AFDC recipients, nearly 63 percent of all terminees were able to obtain unsubsidized employment once leaving a JTPA program.

What then can we make of this information? On the one hand, the data indicate that JTPA does, to a limited extent, differentiate AFDC recipients from others in making service assignments. As we have seen, those receiving AFDC participate in JTPA for a slightly longer period of time, and are more likely to receive training and other special-

---

[1] See for example: Glen Schneider, Michael Battaglia, Christopher Logan, and Jeffrey Zornitsky, An Assessment of Funding Allocation Under the Job Training Partnership Act, Abt Associates Inc., Cambridge, MA, August, 1986.

EXHIBIT 2.3

SELECTED CHARACTERISTICS OF JTPA TERMINEES, BY TYPE OF AFDC RECIPIENT, PY84-PY86

	All Terminees	AFDC Recipients		
		Total	Parents	Female Single Parents
Total Terminees	1,937,844 (100.0)	407,956 (21.1)	274,610 (14.2)	220,109 (11.4)
Education Status				
School Drop-Out	26.5	30.7	33.1	33.2
Student	14.5	16.4	2.5	2.6
High School Grad.	41.6	39.4	47.4	47.2
Post High School	17.3	13.5	17.0	16.9
Labor Force Status				
Employed	10.5	5.1	5.1	5.1
Unemployed	67.8	66.8	75.3	73.9
Out of Labor Force	21.8	28.2	19.5	21.0

Source: Job Training Longitudinal Survey, Unpublished Data

EXHIBIT 2.4

IMMEDIATE PROGRAM OUTCOMES EXPERIENCED BY JTPA TERMINEES, BY TYPE OF AFDC RECIPIENT, PY84-PY86

	All Terminees	AFDC Recipients		
		Total	Parents	Female Single Parents
	1,937,844 (100.0)	407,956 (21.1)	274,610 (14.2)	220,109 (11.4)
Reason for Termination				
Placed in Unsub. Employment	62.7	53.5	58.5	56.9
Return to School/ Training	2.4	2.6	1.7	2.0
Other	34.9	43.8	39.8	41.1
Hourly Wage at Placement	$4.67	$4.47	$4.60	$4.45

Source: Job Training Longitudinal Survey, Unpublished Data

ized services as well as basic education and supportive services. On the other hand, the differences are sufficiently small to suggest that the basic JTPA program structure is short-term in nature regardless of who is served. Moreover, the data on terminee characteristics indicate that most of those served by JTPA -- AFDC recipients and others alike -- are high school graduates with an active attachment to the labor force. Thus, we cannot characterize the JTPA population as severely disadvantaged, at least on the basis of education and labor force status at the time of program entry. Nor can we say that JTPA makes a sharp program distinction between AFDC recipients and other participants.

To some extent, these findings may be an outgrowth of JTPA's principal focus on training and job placement. Absent the mission and resources to provide extensive remediation and support services, the program may well gravitate toward providing services to those for which it can be most effective. As we discuss in the next section, dedication targeting to more disadvantaged groups requires a program mix and funding level far beyond that provided for in the JTPA program.

2.3 Key Program Issues

Although JTPA represents the nation's major employment and training program for AFDC recipients, several states and localities have initiated ambitious efforts to better link public assistance with the full mix of training and related services needed to limit an individual's dependence on public assistance. In most cases, the impetus for these efforts has begun at the state level under the auspices of welfare reform. In many respects, states and localities are now swept up in an effort to better serve the AFDC population and create stronger and more effective links between work and welfare.

In the majority of states in our study, targeting welfare recipients and identifying the most appropriate types of services to provide to them were major policy issues. There was, however, noticeable variation in the degree of sophistication and the stage of development in which states have actually addressed these issues. Within this context, we found that states generally fell into one of two major categories. The first group includes those that are well ensconced in full-fledged efforts to improve the level and quality of services for the welfare population. For the most part, these states have implemented efforts that center on reform initiatives which include specific targeting criteria, provisions requiring coordination between JTPA and departments of public welfare, and enriched program designs that emphasize remediation services as well as more output-oriented services such as occupational skills training.

The second group includes states that are just in the initial stages of broad efforts to improve service delivery and develop targeted approaches to serve the AFDC population. In these states as well, however, most of the attention is aimed at demonstration and reform initiatives, rather than at the ongoing system of existing services.

It is notable that states in both categories are introducing reform efforts in a deliberate and careful manner to protect the existing service delivery systems from destabilization, excess strains on organizational capacity, or other adverse effects of rapid change. Thus, states such as New Jersey and Pennsylvania have chosen to introduce their state-wide welfare initiatives on a gradual basis rather than a universally instantaneous one. This surge in judicious experimentation is a characteristic hallmark of states' efforts to improve employment and training services to welfare recipients.

Despite such promising efforts, several persistent substantive and institutional factors surfaced as constraints to effective service delivery. First, regardless of the status of state reform efforts, we observed a general consensus that the current structure and mix of employment and training programs, by themselves, cannot be expected to resolve the multiple barriers to employment that many AFDC recipients, and especially those who are hard-to-employ, often face. One respondent from a state welfare department told us, for example, that:

> Simply modifying JTPA is not going to solve the whole problem of effectively providing employment and training services to hard-to-employ AFDC recipients. It will help, but it will not be a panacea because the absolute level of JTPA resources is simply not sufficient to serve hard-to-employ clients.

Similarly, our field interviews revealed that SDAs are struggling to develop strategies to cope with the problem that, in the words of one informant, "JTPA cannot be all things to all people." SDAs are experimenting with different approaches to this problem. One type of strategy is to carve out a particular niche from among the JTPA-eligible population on which to focus JTPA service efforts; for example, by making highly targeted efforts to serve severely disadvantaged welfare clients. Yet, SDA directors appeared vitally aware of the implications -- particularly the difficulties -- of such a strategy in terms of increased recruitment and enrollment difficulties, political liabilities, and potential (or actual) loss of (non-targeted) 6 percent funds. According to one SDA respondent:

> "There are lots more attractive opportunities than JTPA on the street. Our training contractors do their own recruiting; they have a rough time because the reputation of the training program is that it is like school. By contrast, our neighborhood-based direct placement

centers have an easier time in doing recruitment because they are viewed as places where you can go and get a job. These recruitment difficulties have been intensifying as our 16-21 age group has been shrinking."

A second type of constraint rests on heightened expectations about what JTPA should be able to achieve. While strong public support for reforming the welfare system has spawned new, innovative state-based efforts, it has also introduced a tension regarding the expected outcome of these initiatives. During our interviews, we noted a conflict that many state and local officials were increasingly faced with as they attempted to improve the level and quality of service delivery to AFDC recipients.

On the one hand, respondents pointed to considerable public concern over whether funds were being spent well. In many cases, this was translated into political pressure to demonstrate on both the state and local service provider level "more bang for the buck." Success in this context was usually interpreted quite narrowly to mean quick and easily quantified results, such as large numbers of low-cost placements. On the other hand, many of the same respondents recognized that serving the AFDC population effectively requires significant, long-term investments of public resources that may not demonstrate visible results in the short run. This perspective was viewed as counter to what many officials believed was necessary to garner public support, and in their view, created an untenable conflict. It nonetheless created clear incentives for officials and program operators to engage in short-term program services that emphasized low costs and immediate placement. For this reason, many states and SDAs viewed consensus building and clarification of the role of work and welfare programs as key to their efforts.

A third type of constraint was suggested when we went into the field. A common theme among our field interviews was the recognition that the current economic climate has created a different set of challenges and opportunities than were in existence when the Congress created the JTPA system in the early 1980s. Specifically, recent growth in the national economy as well as in particular state and regional economies has tightened labor markets such that it is more difficult to enroll those JTPA-eligible participants who only "need a little boost" to obtain stable employment.

In particular, several respondents reported a growing overlap between the characteristics of their JTPA clients and those of the local welfare clientele. One state-level welfare official commented:

"JTPA is philosophically geared towards a different type of individual, i.e., the structurally unemployed individual rather than the welfare

> client with huge up-front deficits in education, life skills, and job readiness. I believe that this partially reflects the economic conditions and concerns that were prevalent when JTPA was being debated and written. While the concern at that time was over high rates of unemployment, we now have full employment in many places and so we need to train the "bottom rung" for employers since everyone else is working. Yet, the JTPA system is not geared to address the problems of this bottom rung."

As would be expected, respondents echoing this sentiment believe that the mission and purpose of JTPA needs to be amended to reflect these new economic realities, to refocus the system towards targeting individuals with multiple barriers to employment, and to allocate the additional resources which will enable SDAs to provide the rich mix of remedial and supportive services that such clients desperately need.

During the field interviews, many of our respondents also indicated that tight labor markets have created severe pressures on existing service delivery patterns. They suggested that highly disadvantaged clients, because of their limited motivation and literacy capabilities, cannot take advantage of the typical, limited mix of JTPA employability services. Yet, they also suggested that JTPA was not designed to fund all the remedial education and other services that such clients need as a necessary step to achieve success within the JTPA system and in the labor market. While most agreed that better coordination with education and social service agencies would help, such arrangements were viewed as difficult to arrange and maintain. Moreover, even under the best of circumstances, SDAs cited the difficulties associated with keeping AFDC recipients in a program long enough to successfully complete training and obtain a job.

As we noted in the interim report, our telephone interviews revealed that state JTPA and welfare department staff universally spoke to the critical role played by supportive services such as child care, transportation, housing allowances, and case management services. Most felt that the existing resources must be enhanced for welfare recipients to become successful program participants and surmount their multiple employability barriers.

Our field interview data strongly confirm these findings. The large majority of field respondents immediately identified limitations on supportive services as key ways in which JTPA limits an SDA's ability to provide services to welfare recipients, particularly the hard-to-employ. As SDAs and service providers try to provide welfare recipients with employment and training services, the limited availability of supportive services often inhibit such efforts in several areas, such as recruitment, enrollment, matching services to individual needs, maintaining participation for the duration of the training,

and achievement of successful outcomes. (Stipends were cited especially critical for particular groups of economically disadvantaged individuals such as those who are not receiving AFDC or homeless individuals who do not have the proper documentation to apply for welfare assistance.)

Our field interviews also reinforced earlier respondents' view that inadequate supportive services are a major factor inhibiting welfare recipients from believing that they can leave the welfare rolls, obtain skills training, and find a decent job without jeopardizing the already-fragile financial well-being of their families. And as many of our respondents indicated, personal motivation and a belief that success is attainable are among the crucial first steps for welfare recipients to undertake in the process of transitioning from welfare to work.

As would be expected, many respondents viewed increased funding as a way to overcome the limited availability of support services. Given the disparate nature of existing programs for the welfare population, such a response is quite predictable. Each agency or program charged with responsibility for serving AFDC recipients naturally views itself as limited by those resources not directly part of its charter. But our respondents also recognized the important role that inter-agency coordination could play in this regard. In theory, most agreed that formal, functional agreements between JTPA, welfare, education, and social service departments could contribute significantly to improving the level and quality of services available to the AFDC population.

In practice, however, coordination was much more difficult to achieve. At both the state and local levels, we observed genuine interest in and efforts to coordinate that were thwarted by institutional and operational problems. Moreover, once coordination agreements had been achieved, they required a substantial amount of constant attention, modification, and political acumen. In one of the SDAs we visited, for example, coordination with the local welfare office was constantly limited by an inability to efficiently share information on client progress. The pressing priorities of each agency, coupled with differences in reporting requirements kept two agencies focused more on the process of coordination than serving the participant effectively.

Despite such turf battles and other barriers to more cooperative agency working relationships, our evidence from the field suggests that progress is being made. As one JTPA informant observed,

> "The most important challenge we face in serving welfare clients and other hard-to-serve individuals is to establish some kind of true collaborative working relationship with the state welfare department. I am

> optimistic about this because some barriers are breaking down. At the local level in particular, turf issues are being dealt with more constructively. The state welfare department knows that they can't really live without JTPA since it is their greatest source of training opportunities. So, we will simply have to learn to work together."

Even if coordination, limited political pressure, and enriched program services could be achieved, our results indicated that other issues would remain to be resolved. First, several respondents revealed some measure of difficulty defining the hard-to-employ and identifying how to best provide new, enriched and comprehensive services for them. The tasks of first identifying hard-to-employ individuals, particularly welfare recipients, and then targeting appropriate employment and training services to them are inextricably related. These two tasks can properly be viewed as both a challenge and a policy vehicle through which to improve welfare recipients' receipt of effective employment and training services.

Establishing formal criteria to identify hard-to-serve individuals was viewed by most of our respondents as a necessary precursor to the development and implementation of targeting and other policies which are intended to encourage SDAs to enroll more at-risk clients. It is notable, however, that relatively little is known or understood about such important topics. Although an explicit focus on targeting hard-to-serve individuals is receiving more attention by states and local agencies, many have yet to formally identify the characteristics of such individuals. Where clarification about such groups is progressing, it is evident that there is substantial variation in the definition of what constitutes the hard-to-employ.

For example, our interviews revealed two distinct perspectives about who the hard-to-serve are within the broad JTPA-eligible population. The first perspective views hard-to-serve individuals, such as welfare recipients, as a distinct and identifiable group within the JTPA-eligible population. This perspective supports policies such as using 6 percent incentive monies in order to encourage SDAs and program operators to recruit and enroll such hard-to-serve clients.

The second perspective was succinctly summarized by one of our SDA respondents:

> "Everyone in JTPA is hard-to-serve; it is simply a matter of degree regarding the barriers to employment that a particular individual is facing."

Additionally, many respondents cited performance standards as another factor limiting the ability of programs to target on the hard-to-employ and provide them with

the full range of services needed to achieve self-sufficiency. This concern was especially evident in the JTPA program, and particularly at the local contractor level where such activity is actually planned and implemented. Four criticisms, in particular, were cited by most of the respondents.

First, many respondents indicated that JTPA's overriding concern for performance achievement had shifted attention away from providing individuals with the services they required, to delivering services in the most efficient manner. While the two are certainly not inherently inconsistent with each other, efficiency was increasingly viewed as synonymous with low costs and high levels of job placement.

In many respects, this perception was shaped by JTPA's cost standards and the relatively short-term nature of all of the standards. Universally, our respondents at the state and local levels viewed the cost standards as limiting the ability of service providers and administrators from providing longer-term, comprehensive services to the AFDC population, particularly those identified as hard-to-employ. As a result, we observed that some state JTPA offices have dropped the cost standard or are limiting its importance in the distribution of 6 percent incentive funds.

While most respondents agreed that the Department of Labor's adjustment models represented a tool for addressing these problems, they also pointed out that the models provided little incentive to change behavior. The relatively small size of the adjustment factors and the exclusion of many variables believed to affect employability (e.g. motivation, literacy) were frequently cited as limiting the effectiveness of the adjustment approach. Moreover, decisions to target on the hard-to-employ appeared to be made independent of the standards and the availability of modeled adjustments. By design, JTPA's performance standards have been set at levels that the vast majority of SDAs are expected to be able to meet. And indeed, the SDAs we visited indicated only limited concern over their ability to achieve most of their performance expectations.

Finally, our JTPA respondents clearly indicated that the performance management system lacked sufficient incentives and technical assistance to encourage the development and delivery of effective services. While the states we visited were increasingly focusing their 6 percent funds on target group requirements, the SDAs did not always view the size of the incentives as sufficient to offset the risks of serving the hard-to-employ. And even in those cases where such risks were being taken, there was a unanimous opinion that more technical assistance on program development and coordination was needed. Simply stated, employment and training programs appear to need more assistance overcoming the barriers to successfully serving the AFDC population.

It would thus appear that creating effective program responses for the welfare population involves several complex programmatic and institutional factors. Clarity over targeting expectations, flexible and comprehensive program services through inter-agency coordination, and the availability of adequate levels of supportive services are, for example, among some of the more important ingredients of effectively operating programs. Efforts to improve the responsiveness of the employment and training system will thus have to consider these factors as well as modifications to the performance standards system.

We now turn to a brief summary of emerging state policies and programs designed to more effectively serve welfare recipients.[1]

2.4 Summary of State Effort to Serve AFDC Recipients

States are now engaged in substantial activity aimed at developing new policies and programs designed to strengthen service delivery to AFDC recipients and to target such services on those identified as hard-to-employ. We observed noticeable interstate variations in the level, types, and degree of sophistication of these efforts that can be explained by varying degrees of political pressure, prior experience with demonstration initiatives, the degree of flexibility in the existing service delivery system, and the ability to clearly identify the characteristics of hard-to-employ individuals. Nonetheless, several basic program elements appear to speak directly to the key problems and constraints discussed earlier in the chapter.

While several states are relying upon their performance standards system to encourage service delivery to AFDC recipients and especially those identified as hard-to-employ, perhaps the most significant vehicle used by states to expand and improve the delivery of services are state-wide welfare reform and other pilot demonstration programs. Key characteristics of these new programs are summarized in Exhibit 2.5.

As can be seen, regardless of the scope of each effort, their key goals are fundamentally similar, although not identical. First, there is a recognition that hard-to-serve clients need an "enriched" mixture of program services which address both their

---

[1] A fuller discussion of these policies and programs is presented in our interim report. See: Mary Rubin and Jeffrey Zornitsky, Serving AFDC Recipients: Initial Findings on the Role of Performance Standards, National Commission for Employment Policy, Washington, D.C., May 1988.

EXHIBIT 2.5

KEY PROGRAM CHARACTERISTICS OF STATE-LEVEL EFFORTS

SCALE OF PROJECT

- Broad-scale initiatives:[1] New Jersey, Pennsylvania, Illinois, Florida, Michigan, Massachusetts, New York
- Smaller-scale Pilot Projects: Texas, Ohio

PROJECT OBJECTIVES

- Extensive Reform: New Jersey, Pennsylvania, Massachusetts, Florida
- Incremental Change: Texas, Ohio

FORMAL INTER-AGENCY COORDINATION

Pennsylvania, Illinois, Florida, Texas, Michigan, New York

ELEMENTS OF ENRICHED PROGRAM MIX

- Case Management: Massachusetts, Florida, Pennsylvania, New Jersey, Illinois, New York
- Enriched Supportive Services: New Jersey, Pennsylvania, Illinois, Florida, Michigan, Massachusetts, New York
- Post-Placement Transitional Health and Child Care Benefits: New Jersey, Pennsylvania
- Longer-Term/Multiple Services: Michigan, Pennsylvania, New York

SPECIFIC TARGETING ON HARD-TO-SERVE CLIENTS[2]

Michigan, New Jersey, Illinois, Pennsylvania, New York, Texas, Ohio, Massachusetts, Florida

PREDOMINANT USE OF PERFORMANCE STANDARDS AND/OR PERFORMANCE-BASED CONTRACTING

Massachusetts, Florida, Pennsylvania, New York, Michigan

[1]May also include smaller-scale pilot projects.

[2]See Exhibit 2.6 for specific targeting characteristics.

EXHIBIT 2.5 (continued)

KEY PROGRAM CHARACTERISTICS OF STATE-LEVEL EFFORTS

MONITORING OF SERVICE PATTERNS AND OUTCOMES

Pennsylvania, Michigan, New Jersey, Massachusetts, New York, Illinois, Florida, Ohio

NEW STATE AUTHORIZING LEGISLATION

Pennsylvania, Florida, Michigan

EXTENSIVE FLEXIBILITY IN PROGRAM DESIGN AND SERVICE DELIVERY SYSTEM

Massachusetts, New Jersey, Pennsylvania, Michigan, New York

lack of basic skills and their need for vocationally-specific skills. Second, there is an expectation that outcomes for this group will improve with a more intensive concentration of services. Third, there is a strong emphasis on inter-agency coordination at both the state and local levels.

At the state level, the demonstrations typically involve the state department of welfare, JTPA, social services, and in some cases, the state department of education. These agencies have worked together to identify respective roles and responsibilities, to develop efficient means of making their programmatic resources available, and to define procedures for directly monitoring program activities. In several cases, these groups have come together as part of top-level inter-agency task forces created to address issues related to welfare dependency. In other cases, they have been brought together as a result of the demonstration projects and remain involved through that vehicle. Typically, the state welfare department has assumed the lead role.

At the local level, the program design of the demonstrations also encourages cooperation between public agencies such as county welfare offices, SDAs, local Employment Service offices, and local education agencies. The basic rationale here is to streamline client flow between these public agencies, to facilitate welfare recipients' receipt of basic services, such as employment and training, public assistance, and education, and to prevent them from becoming discouraged as they work towards economically self-sufficiency. This is expected to yield a more effective allocation of already-available training and educational program slots to those welfare recipients and other hard-to-serve individuals who could most benefit from such opportunities. Thus, the intent of institutionalizing such cooperation is to avoid the creation of a wholly new employment and training system for welfare recipients. A further rationale is to avoid duplication of effort in such activities as eligibility determination.

The service delivery approaches adopted in these initiatives have no particular uniformity across states or between demonstration projects within a single state. Flexibility, innovation, case management services, and an enriched program mix are, however, typical hallmarks. Our respondents observed that flexibility in program design is crucial to allow for the adaptation of basic program design elements to address the programmatic challenges created by particular local conditions or clientele.

Typically, the authorizing legislation and/or policy directive (e.g., via an RFP) sets the core features of a program but allows a broad level of discretion in setting the range, mix, and level of program services. This emphasis on a flexible program mix is also reflected in the increasing reliance on a case management approach. Case

management services are designed to prepare and coordinate comprehensive employment plans for participants, to assure access for the necessary training and support services, and to provide support and counseling during program participation and also after job placement.

Several of our respondents noted that this enhanced client/staff contact was designed to counter the inability of traditional welfare-related employment and training programs to provide more than minimal personalized staff attention. One of our welfare department respondents observed:

> A typical JTPA program doesn't do case management, so case management will be a key component in helping to achieve the performance measures in [new state welfare reform initiative] by, for example, helping to keep dropouts low. That is, hard-to-serve welfare clients have high dropout participation rates in traditional employment and training programs. So, case management as well as supportive services will help these participants stay in the program and achieve success.

Across all the major demonstration projects, we noted requirements that support services, such as child care and transportation, be provided in sufficient quantity as core services during program participation. These are viewed as crucial to the ability of welfare recipients, particularly hard-to-serve subgroups of the AFDC population such as young mothers with children under 6 years of age, to enroll in and successfully complete a comprehensive set of program components. Notably, many demonstrations continue to provide supportive services for up to one year after job placement as well.

We also observed that adequate and stable funding was viewed as allowing innovative planning without uncertain or narrow resource constraints. Funding was usually a mixture of currently available federal funds for public assistance programs (such as WIN and the Food Stamp program) and for employment and training programs (such as JTPA and the Employment Service) together with special legislative allocations of state funds.

Finally, in most cases the demonstration and reform initiatives included specific performance standards. For the most part, these standards were included in performance-based contracts and included payment provisions that were tied to the achievement of specific benchmarks, such as job placements and post-program retention. Overall, our respondents uniformly believed that including performance standards in service contracts was critical to the success of their efforts.

Just as these demonstration programs emphasize innovation and comprehensive-

ness, they are also targeted on selective segments of the AFDC population believed to reflect the hard-to-employ population. Exhibit 2.6 presents the specific targeting provisions contained in the welfare reform initiatives implemented in the states we surveyed. Several notable features emerge. First, there is noticeable diversity among the states with respect to the characteristics of the hard-to-employ.

Second, the emphasis of these targeting policies varies between groups that are most at risk of becoming long-term welfare dependent (e.g., teen mothers) and groups that are thought to have low employability potential (e.g., high school dropouts, the homeless, individuals with severe literacy problems). Although states do recognize that all groups which are the focus of targeting efforts typically share common barriers to employment, such as lack of specific vocational skills or prior work history, the differences between their targeting policies reflect the special barriers to employment faced by a particular group. Overall, while we observed diversity among the targeting policies, it is important to point out that the majority of states' policies are targeted toward female welfare recipients who are exempt from mandatory WIN participation because their children are under age six.

If we take these initiatives as a reasonable reflection of the types of program models required to serve the welfare population effectively, then they would suggest that simply changing JTPA's performance standards will not necessarily lead to large changes in the program's targeting priorities and program mix. It would appear that other, more basic changes in program design and targeting requirements would also be necessary to forge a coordinated and more comprehensive service delivery approach.

EXHIBIT 2.6

CHARACTERISTICS OF HARD-TO-SERVE TARGET GROUPS IN STATE-BASED WELFARE REFORM INITIATIVES

State	Program Name (Type)	Hard-to-Serve Target Groups
Pennsylvania	Single Point of Contact (SPOC) Demonstration	Single parents with children under age 6
		Single parents who have been on public assistance for at least two years
		Individuals with less than a 6th grade education
		Individuals with severe literacy problems
	Transitionally Needy (TN) Demonstration	Employable recipients of General Assistance who have multiple barriers to employment
New York	Comprehensive Employment Opportunity Support Centers (CEOSC) Program (state-wide)	Public assistance recipients with children under age 6
		Current or former teen parent who is on public assistance
		An individual who has been on the welfare rolls for two years or longer (Note: New York recently changed its definition of long-term receipt from four years to two years, in response to recent research on the dynamics of welfare receipt.)
Texas	Unnamed program	16-17 year old dropouts who are mandatory (WIN) registrants
		Newly certified welfare recipients
		Individuals who are received AFDC for 2+ years
		Individuals who have received AFDC for 2 years or less

EXHIBIT 2.6 (continued)

CHARACTERISTICS OF HARD-TO-SERVE TARGET GROUPS

IN STATE-BASED WELFARE REFORM INITIATIVES

State	Program Name (Type)	Hard-to-Serve Target Groups
Michigan	Michigan Opportunity and Skill Training (MOST) Program (statewide)	Primary wage earner in an AFDC-UP household (although the spouse can be a voluntary MOST registrant)
		Teens who are age 16 or older, not in school full-time, and part of an AFDC case
		An individual who is economically disadvantaged
		An individual who has been unemployed 2+ years or never been employed
		An individual who is in need of functional literacy or employability characteristics or supportive services
		An individual whose employability development plan states that the length of time or cost of training leading to employment exceeds that received by the average person enrolled in a designated area
Ohio	Supported Work Demonstration	Individuals who have been on public assistance for 3 to 6 years
		Parents aged 18 to 21 with children under age 6
		Handicapped individuals
Massachusetts	ET Choices (statewide)	Individuals receiving AFDC for 2+ years
		Teen parents
		Hispanics
		Public housing residents

EXHIBIT 2.6 (continued)

CHARACTERISTICS OF HARD-TO-SERVE TARGET GROUPS

IN STATE-BASED WELFARE REFORM INITIATIVES

State	Program Name (Type)	Hard-to-Serve Target Groups
Florida	Project Independence (state-wide)	Female AFDC recipients with children age 3 and older
Illinois	Project Chance (state-wide)	Female AFDC recipients with children over age 6
		Male General Assistance recipients in Chicago
		AFDC recipients in public housing
		Parents under age 21
New Jersey	Realizing Economic Achievement (REACH) Program (state-wide)	Welfare applicants with children age 2 or older
		Welfare recipients with children age 2 or older

3.0 KEY ISSUES IN DEVELOPING A PERFORMANCE STANDARDS SYSTEM

In this chapter, we present the performance standards system implemented under JTPA and highlight major issues that have arisen in its use. We then present a critical discussion of key challenges in designing a performance management system for programs serving AFDC recipients and other hard-to-employ groups.

3.1 JTPA's Performance Standards System[1]

Performance standards have long been regarded as an important management tool for employment and training programs. They provide basic information for addressing key strategic planning and operational decisions, and allow funding agencies to determine the extent to which programs are moving in a direction consistent with their long-term goals. For the manager who operates within a dynamic, rapidly changing environment, performance standards provide a timely and useful indication of program effectiveness.

The art of managing with performance standards has been developed largely on the basis of experience gained under employment and training programs authorized first by the Comprehensive Employment and Training Act (CETA) and currently by the Job Training Partnership Act (JTPA). While employment and training programs authorized by other sources have also relied upon performance standards, those administered by the Department of Labor have assumed a leadership position in this area of program management.

In many respects, the implementation of JTPA in October 1983 continued the growing emphasis that was placed on performance management in employment and training programs. However, unlike CETA, JTPA was crafted in the image of a performance-driven program with a much more elaborate management structure designed to clearly link program inputs and outputs. Moreover, the crafters of the legislation included very specific guidelines regarding the choice of performance measures, the basis upon which to develop numeric standards, and rewards and sanctions for exceeding and consistently falling below performance expectations. And, of equal

---

[1]For a complete discussion of the evolution of performance standards, see: Rubin, Mary, Zornitsky, Jeffrey, Serving AFDC Recipients: Initial Findings on the Role of Performance Standards, National Commission for Employment Policy, Washington, D.C., May, 1988.

importance was the transfer of responsibility for managing performance standards from the Department of Labor's regional offices to state government. In many respects, the performance standards system created for JTPA has been characterized as an effort to introduce market-like forces and incentives into federal employment and training programs.[1]

Like CETA, the performance measures established under JTPA are intended to be reasonable proxies for long-term net impacts, and when used in an objective manner, are expected to promote the Act's overall goals. Currently, JTPA operates with seven termination-based performance standards; three thirteen-week post-program measures have been identified as candidates for Program Year 1988. These measures and their national averages are presented below.

JTPA Title IIA Performance Standards
(PY 1986/87)

Performance Measure	National Average
Adult entered employment rate (AEER)	62.00%
Adult cost per entered employment (ACEE)	$4374.00
Adult average wage at placement (AAWP)	$4.91
Adult welfare entered employment rate (AWEER)	51.00%
Youth entered employment rate (YEER)	43.00%
Youth positive termination rate (YPTR)	75.00%
Youth cost per positive termination (YCPT)	$4900.00
Adult post-program employment rate	N/A
Adult average post-program weekly earnings	N/A
Adult average post-program weeks worked	N/A

Consistent with the intent of Congress, much of the responsibility for managing performance standards is placed at the state level, although the Department of Labor retains certain important duties. For programs funded under Title II of JTPA, the basic federal role is defined in Section 106 as follows:

> (b)(1) ...for adult training programs under Title II... the Secretary shall prescribe standards on the basis of appropriate factors which may include (A) placement in unsubsidized employment, (B) retention in unsubsidized employment, (C) the increase in earnings, including hourly wages, and (D) reduction in the number of individuals and families receiving cash welfare payments and the amounts of such payments.

---

[1]See Thomas Bailey, "Market Forces and Private Sector Processes in Government Policy: The Job Training Partnership Act, in Journal of Policy Analysis and Management, Vol. 7, No. 2, 1988, pp. 300-315.

(2) ...the Secretary shall also designate factors for evaluating the performance of youth programs which...shall be (A) attainment of employment competencies recognized by the private industry council, (B) elementary, secondary, and postsecondary school completion, or the equivalent thereof, and (C) enrollment in other training programs or apprenticeships, or enlistment in the Armed Forces.

(4) The Secretary shall prescribe performance standards relating gross program expenditures to various performance measures.

The standards must also include provisions governing the base period prior to participation, a representative period after termination, and cost-effective methods for data collection.

The authority of the Federal government to prescribe utilization of the standards is, however, clearly circumscribed by the Act. For example:

(e) Each Governor may prescribe, within parameters established by the Secretary, variations in the standards...based upon specific economic, geographic, and demographic factors in the State and in service delivery areas within the State, the characteristics of the population to be served, and the type of services to be provided. (JTPA, Section 106(e))

To further support its emphasis on performance management, the Act also places performance standards prominently in the state and local planning process, and provides rewards and sanctions that are directly linked to local SDA performance. In the first case, Section 104 requires SDAs to include in their annual plans performance goals as well as plans for the preparation of an annual report that must address the extent to which these goals were achieved. In addition, Section 105(b)(1) requires Governors to approve SDA plans unless corrective measures for not meeting performance expectations have not been taken or are not underway.

In the second case, Section 106(h) authorizes the Governor to provide technical assistance to low performing SDAs and to apply sanctions if poor performance persists for at least two consecutive years. Further, Section 202(a)(3) makes available six percent of the state allocation for distribution to those SDAs exceeding their performance standards, for technical assistance, and for services to hard-to-employ individuals.

To fulfill the Secretary of Labor's responsibilities under JTPA, the Employment and Training Administration implemented regulations and a suggested framework for Governors to use for establishing performance standards. One of the most noteworthy features of this framework is the optional adjustment methodology for setting local performance expectations. This methodology builds upon the CETA experience by using regression techniques to account for factors out of the control of local service providers

when setting standards. Since its initial implementation, however, there has been growing emphasis placed on encouraging states to go beyond the model in setting local performance standards by considering those factors which may affect performance but are not included in the standard adjustment. In many respects, the current trend in JTPA performance management is headed toward the general approach used in the later stages of CETA which emphasized more of a balance between statistical forecasting and negotiation over key factors that influence participant employability.

Although JTPA's performance standards system is expected to fulfill a critical management role, it can also create "perverse incentives" to serve the most job-ready applicants in short-term, low cost programs. To counteract these potential effects, JTPA contains at least five mechanisms, including:

1. A requirement that 90 percent of all enrollees be disadvantaged;
2. Targeting provisions requiring specified levels of service to youth, high school drop-outs, and welfare recipients;
3. Authorization for the Governor to adjust SDA standards to account for variations in participant characteristics and other factors affecting performance;
4. The establishment of performance standards at a minimally acceptable level; and
5. The availability of financial incentives that Governors can use to influence SDA behavior.

Despite these efforts, it is possible that performance standards may have the effect of reducing services to hard-to-employ individuals or limiting the investments that are made in participants. Indeed, this has been argued by practitioners, researchers, and representatives from the Congress. Many believe that the performance standards system has shifted the attention of programs away from long-term employability development to the creation and management of a low cost, high volume service delivery system.

Whether performance standards cause "creaming" and/or the delivery of low-cost services is one of the most controversial and confusing issues surrounding JTPA. While evaluative evidence on this issue is not yet available, at least four factors

contribute to the concern.[1] The first is widespread use of screening criteria for admission into JTPA programs. Program operators have been often found to screen out the most disadvantaged, through education and skill tests, in order to meet the performance standard provisions of their contracts. The second factor is the relatively limited investment that JTPA makes in its participants. For example, between July 1984 and June 1987, the duration of service receipt was relatively short, even for the more expensive training services and the more difficult-to-serve participants. Overall, the median length of stay for Title IIA terminees was approximately 15 weeks; in job search assistance it was less than four weeks, while in classroom training, the median was roughly 19 weeks.

The third factor is the relatively low service share received by high school dropouts. In contrast to a 51 percent share of the eligible JTPA population, service shares to this target group were 37 percent in PY84 and 40 percent in PY85. Although these shares closely conform to JTPA's targeting provisions, they can still be viewed as below the relative service needs of high school dropouts.

The fourth and final factor is growing concern voiced by the practitioner community that the performance standards system is simply constraining local discretion over whom to serve and what types of programs to offer. Four particular concerns about the performance standards system are advanced, including:

1. The statistical adjustment models do not fully neutralize the risks associated with serving participant subgroups, and particularly AFDC recipients and other hard-to-employ groups. While the adjustment models do control for several important characteristics, they do not account for intra-group differences in characteristics that affect employability nor do they control for other key characteristics that also influence performance achievement, such as literacy and motivation. Thus, program operators argue that they remain confronted with disincentives to serve hard-to-employ groups both within and across those already included in the adjustment procedure.

---

[1]Support for many of these concerns can be found in: Westat, Inc., Transition Year Implementation of the Job Training Partnership Act, January 1985; Grinker Associates, Inc., An Independent Sector Assessment of the Job Training Partnership Act, Grinker Associates, Inc., New York, July 1986; Gary Orfield and Helene Slessarev, Job Training Under the New Federalism, Unemployment and Job Training Research Project, Chicago, Illinois, 1986; U.S. General Accounting Office, The Job Training Partnership Act: An Analysis of Support Cost Limits and Participant Characteristics, GAO, Washington, D.C., November 6, 1985; and U.S. Department of Labor, Summary of JTLS Data or JTPA Title IIA and III Enrollment and Terminations During July-December 1985, Washington, D.C., May 1986.

2. Although the Department of Labor has developed a number of additional procedures for adjusting performance beyond the model, anecdotal evidence suggests that many states may not be utilizing them, and in those cases where they have been implemented, SDAs do not readily take advantage of them. While the reasons for this are unclear, many believe that they stem from the perceived complexity of the standards system generally, and the data and other requirements associated with these additional adjustment procedures in particular.

3. The standards themselves and particularly the cost standards present a clear message that costs must be limited, thus constraining the type of investments that service providers can make and the types of individuals they can enroll.

4. JTPA's overall emphasis on performance achievement has created a culture that places undue priority on the performance standards at the expense of other program goals such as serving the hard-to-employ.

These developments are not evidence of the effects of performance standards per se; in fact, a study recently released by the National Commission for Employment Policy reports that with the exception of adult high school dropouts, other target groups including AFDC recipients, have been served in a manner proportional to their representation in the eligible population.[1] However, when combined with JTPA's overriding emphasis on performance standards, these developments have raised "red flags" and lead many to consider whether the program's performance management system needs to provide greater stimulus for serving the hard-to-employ.

At the same time, performance standards cannot be viewed in a vacuum. Other aspects of JTPA may also be contributing to concerns over creaming and limited program investments. Perhaps most important is that JTPA is not targeted on the hard-to-employ. Eligibility criteria for Title IIA programs are income-tested, focused almost exclusively on the economically disadvantaged, and sufficiently broad to permit local program operators to provide services to those who, in their judgment, can benefit from and are in most need of services. Thus, JTPA has a relatively large eligibility pool from which it must select economically disadvantaged individuals who are most in need of services and who can also benefit from service receipt.

---

[1] See: Steven Sandell and Kalman Rupp, op. cit.

Related to this issue is the absence of unanimity regarding the characteristics of the hard-to-employ population. To date, neither the Department of Labor nor the Congress have explicitly defined the hard-to-employ population, making it difficult to both judge existing service patterns and provide direction over whom to serve in the future. And even if the hard-to-employ were defined, it would remain to select those who could benefit most from program services. Absent a national definition, the very structure of JTPA would encourage variation on both a state and local level.

In addition to targeting flexibility, JTPA's limitation on the use of stipends and supportive services places greater reliance on individual self-selection among program eligibles. Participant motivation thus takes on a prominent role in JTPA and may well run counter to the desire of providing services to those with relatively limited incentives to seek work or assistance finding it.

Like performance standards, the relative importance of these factors remains untested. They do, however, suggest that performance standards alone may not be at the cutting edge of the creaming and low-cost service problem. Modifying the performance standards system may thus be a necessary but not a sufficient condition for improving the level and quality of services to the hard-to-employ.

Nonetheless, the seriousness of the issue has caused the Congress and the Department of Labor to reassess the performance standards system critically and to consider modifications to it. At the same time, the Congress and the Administration are also considering amendments to JTPA that would encourage services to AFDC recipients and other hard-to-employ groups, as well as legislation that would restructure the welfare system and the manner in which training services are provided to public assistance recipients. Since several of the legislative proposals suggest the creation of performance standards to guide the provision of services to welfare recipients, an assessment of such standards can provide timely input to current policy deliberations.

3.2 Developing a Performance Management System

Modifying the performance standards system to further support the goal of serving AFDC recipients raises a number of substantive issues. While these issues are not new in a conceptual or research sense, their application to the hard-to-employ presents new challenges.

The first challenge is developing a set of performance measures that are valid predictors--in both a statistical and intuitive sense--of JTPA's objectives. Although

substantial research has already been conducted on the validity of performance measures for employment and training programs, there are reasons for questioning their application to training programs serving welfare recipients. First, each of the prior studies used as validation benchmarks impact estimates derived from quasi-experimental methods.[1] Since it has been demonstrated that such estimation techniques yield biased results of an often unknown magnitude and direction, it would be difficult to defend the use of such results in a new welfare reform program or in an effort to modify the existing JTPA performance measures. To the extent feasible, it would be desirable to reassess the existing measures using experimental data.

Even if existing validation findings were substantiated using experimental impact estimates, it would also be necessary to determine if the indicators are also appropriate for programs targeted on welfare recipients. Since prior research has been based on the economically disadvantaged population, as opposed to welfare recipients, and because these studies have used earnings, not welfare dependency, as the benchmark measure, there is no a priori reason to assume that their findings can be fully generalized to a targeted segment of the disadvantaged population. In fact, recent evidence would suggest a potential problem and the need for further investigation.

To see this, consider the information presented in the following three exhibits. Exhibits 3.1 and 3.2 display the most recent validation findings for employment and training programs, using earnings as the benchmark. As can be seen, nearly every measure tested is significant and has the expected positive sign, suggesting that JTPA's current performance indicators are valid management tools. Given their predictive

---

[1]Each of the following validation studies relied upon quasi-experimental methods to derive benchmarks:

Michael E. Borus, "Indicators of CETA Performance," Industrial and Labor Relations Review, Vol. 32, No. 1, October 1978.

Katherine Dickinson, Terry Johnson and Richard West, An Analysis of the Impact of CETA Programs on Components of Earnings and Other Outcomes, SRI International, Menlo Park, CA, June 1984.

Robert Gay and Michael E. Borus, "Validating Performance Indicators for Employment and Training Programs," The Journal of Human Resources, Vol. 15, No. 1, winter 1980.

V.J. Geraci and C.T. King, Employment and Training Program Performance: Long-Term Effects and Short-Term Indicators, Center for the Study of Human Resources, University of Texas at Austin, TX, September 1981.

Jeffrey Zornitsky, et al., Establishing a Post-Program Performance Management System for Employment and Training Programs, Abt Associates Inc., Cambridge, MA, 1986.

Exhibit 3.1

Title IIA Validation Correlations for Adjusted Level Measures Using CLMS Data

	TERMINATION			3-MONTH		
Performance Measures	Net Impact I (Fixed Effect Model)	Net Impact II (Matched Comparison Group)	(Gross Adjusted Earnings)	Net Impact I (Fixed Effect Model)	Net Impact II (Matched Comparison Group)	(Gross Adjusted Earnings)
Employed or not	.079***	.131***	.210***	.115***	.195***	.273***
Weeks employed				.132***	.218***	285***
Hours worked				.141***	.230***	.316***
In labor force or not				.073***	.144***	.175***
Weeks in labor force				.080***	.144***	.171***
Wage at end of period	.062***	.155***	.245***	.109***	.234***	.315***
Wage during period				.089***	.220***	.292***
Total earnings				.120***	.264***	.348***
Weekly earnings (Employed only)				.031	.210***	.205***

*Significant at 10%.

**Significant at 5%.

***Significant at 1%.

SOURCE: Zornitsky, Jeffrey, et al, Establishing a Post-Program Performance Management System for Employment and Training Programs, Abt Associates Inc., Cambridge, MA, December 1986.

Exhibit 3.2

Validation Correlations for Adjusted Title IIA Change and Level Performance Measures Using CLMS Data

Performance Measures	3-MONTH			6-MONTH		
	Net Impact I (Fixed Effect Model)	Net Impact II (Matched Comparison Group)	(Gross Adjusted Earnings)	Net Impact I (Fixed Effect Model)	Net Impact II (Matched Comparison Group)	(Gross Adjusted Earnings)
Change Measures						
Weeks employed	.103***	.182***	.198***	.108***	.193***	.197***
Hours worked	.072***	.106***	.121***	.078***	.117***	.117***
Weeks in labor force	.108***	.130***	.135***	.118***	.136***	.140***
Wage during period	.044***	.100***	.090***	.062***	.116***	.080***
Total earnings	.136***	.225***	.275***	.147***	.239***	.272***
Level Measures						
Weeks employed	.132***	.218***	.285***	.160***	.283***	.356***
Hours worked	.141***	.230***	.316***	.159***	.281***	.371***
Weeks in labor force	.080***	.144***	.171***	.108***	.203***	.230***
Wage during period	.089***	.220***	.292***	.108***	.268***	.334***
Total earnings	.130***	.264***	.348***	.154***	.339***	.422

*Significant at 10%.

**Significant at 5%.

***Significant at 1%.

SOURCE: Zornitsky, Jeffrey, et al, Establishing a Post-Program Performance Management System for Employment and Training Programs, Abt Associates Inc., Cambridge, MA, December 1986.

validity, one could assume that they are appropriate for targeted welfare programs. However when we consider the findings in the next exhibit, such an assumption becomes tenuous.

Exhibit 3.3 presents recent validation findings obtained by the Manpower Demonstration Research Corporation (MDRC) based on data collected from their work-welfare demonstrations. Using experimental earnings and welfare savings impacts, the researchers attempted to validate whether employment or welfare status soon after program termination were good predictors. The results are disappointing and suggest that at least two intuitively appealing performance measures may be inappropriate for targeted welfare programs. As the study summarized:

> This conclusion--which runs counter to common wisdom--simply reflects the fact that the magnitude of the program effect on finding a job or leaving welfare is greater for some groups of individuals than others. This does not imply that programs should stop trying to help all people in the caseload find jobs and leave welfare. It does mean that judging programs on the basis of these outcome measures--without considering differences in caseload characteristics and economic conditions--is unwise. It is quite possible, for example, for a program with a relatively low placement rate in a poor labor market to have greater impacts than another program with a more job-ready caseload and more placements. The analysis also shows that this conclusion does not change when longer-term employment rates are substituted for immediate job entries.[1]

While these findings were obtained from a particular demonstration program, they do cast some doubt on the use of gross, short-term performance measures. At a minimum, the results suggest that more validation research, using experimental data from training programs targeted on welfare recipients is quite desirable. The findings also imply that welfare-related performance standards should be adjusted for the background characteristics of individuals. While it can be argued that the existing adjustment model accomplishes this, the current adjustments do not account for at least one key characteristic which improved predictive validity in MDRC's research--that being prior work history. Recall that MDRC found that the predictive validity of the two candidate measures improved substantially when they were adjusted for earnings during the year prior to program participation. Thus, even the existing welfare adjustment model may be

---

[1]Daniel Friedlander and David Long, A Study of Performance Measures and Sub-Group Impacts in Three Welfare Employment Programs, Manpower Demonstration Research Corporation, New York, N.Y., March 1987, pp. xii.

Exhibit 3.3

Recent Findings on the Validity of Candidate Performance Measures for AFDC Recipients Receiving Employment and Training Services

Program and Welfare Status	Validity of Job Entry		Validity of Welfare Status	
	Earnings Gain	Welfare Savings	Earnings Gain	Welfare Savings
San Diego, Applicants	poor	weak	fair	weak
Baltimore, Applicants	poor	poor (fair)	poor	poor
Baltimore, Recipients	poor	weak	fair	good
Baltimore, All AFDC	weak	weak (fair)	fair	fair

NOTES: This table summarizes the correlations between the designated indicator and the earnings gains or welfare savings. The following symbols are used:

- Good, indicates a correlation that has the correct sign and is statistically significant.
- Fair, indicates a correlation that has the correct sign but is not statistically significant.
- Weak, indicates a correlation that has the wrong sign but is not statistically significant.
- Poor, indicates a correlation that has the wrong sign and is statistically significant.

SOURCE: Daniel Friedlander and David Long, A Study of Performance Measures and Sub-Group Impacts in Three Welfare Employment Programs, Manpower Demonstration Research Corporation, New York, N.Y., March 1987.

inadequate as currently specified to reduce creaming incentives within the AFDC population substantially.

It would thus appear that indeed, more research is needed to identify valid performance measures for targeted welfare programs. While research to date provides many consistent findings, it remains to test their robustness with experimentally derived impacts of welfare reductions for public assistance recipients served in employment and training programs.

The second major challenge is to design a methodology and management strategy that can be used to set performance standards. Such a methodology needs to not only address statistical forecasting criteria, but it must also be sufficiently informative, practical, and flexible to maximize the targeting of services to recipients, and especially to hard-to-employ individuals. As discussed above, the MDRC findings suggest that the existing model for adjusting the adult welfare entered employment rate may be incomplete with respect to background characteristics. If this is the case, then the adjusted standards may be biased upward, at least in terms of serving long-term, hard-to-employ public assistance recipients. Unless additional data collection requirements are forthcoming, new ways for adjusting standards may well be required.

Establishing a flexible and informative performance standards setting process needs to go beyond a complete, valid adjustment model. Funding agents and program operators need assistance and information to determine how to set their standards and take account of difficult-to-measure obstacles faced by prospective participants. The best of models cannot be expected to include all important explanatory variables, due to both measurement and data collection limitations. Thus, the standards setting process needs to be viewed in a collaborative context, where funding agents, such as the Department of Labor and states, provide technical assistance and information in order to inform the local decision-making process.

4.0 SELECTING VALID PERFORMANCE MEASURES

The first critical step in developing a flexible and informative performance management system is selecting performance measures that are valid predictors of the key goals of JTPA as well as other employment and training programs serving AFDC recipients. In this study, we assessed the validity of a number of alternative performance measures in predicting two key program goals, net earnings gains and net reductions in public welfare benefits received.

The basic approach we rely on consists of correlating the performance measures being tested (e.g. placement status, post-program earnings) with a second measure that is assumed to be a valid reflection of the underlying goals to be measured (e.g. earnings gains and reductions in welfare dependency), hereafter referred to as the validation benchmark. The stronger the correlation between the test measures and the validation benchmark, the greater the predictive validity of the performance measure.

Ultimately, the quality of predictive validation tests is dependent on the quality of the validation benchmark. As discussed in the previous chapter, a major short-coming of previous studies of performance measure validity is their reliance on quasi-experimental techniques to compute estimates of the validation benchmark. For this study, we rely upon experimental data made available from the AFDC Homemaker-Home Health Aide demonstrations to estimate the benchmarks and conduct the validation tests. While limited in scope, these data provide an opportunity to assess the merits of alternative performance measures for a targeted welfare program operated under experimental conditions.

Predictive validity, however, is not the sole criterion for selecting performance measures. When faced with alternative valid indicators, one must also consider the feasibility and cost of obtaining the data needed to measure the indicator. While, for example, longer-term post-program indicators may be superior to those measured at program termination or shortly thereafter, they also require more lead time for use by management. Similarly, earnings or welfare benefit measures may be more valid predictors than simple status indicators, such as employment status, but they are more complex to estimate. As a result, the final selection of candidate performance measures must be based not only on statistical considerations, but also on qualitative judgements

---

*Section 4.2 of this chapter was prepared by Dr. Stephen H. Bell.

regarding administratively feasibility and cost.

In this chapter, we present the candidate performance measures selected for testing, the procedures used to estimate the two validation benchmarks, and the results that emerged from our assessment of the relative predictive validity of each of the measures. We then consider the issue of administrative feasibility and cost.

4.1 Performance Measures to be Validated

The JTPA program and the welfare reform initiatives introduced by the Congress both provide for a substantial amount of flexibility in the selection of performance measures. Although the range of potential measures varies substantially, the principle focus remains on relatively short-term measures of employment, earnings, and welfare dependency status. On the basis of discussions with NCEP staff and limitations imposed by the available data set, we narrowed the focus of our validation tests to the five basic groups of performance measures presented below.

CANDIDATE PERFORMANCE MEASURES TESTED

Measure	Time: At Termination	Time: 3 Months Post-Termination	Time: 6 Months Post-Termination
1. Employment Status	x	x	x
2. Average Hourly Wages	x	-	-
3. Average Weekly Earnings	-	x	x
4. Welfare Status	-	x	x
5. Total Welfare Benefit Levels	-	x	x

The employment status measures reflect whether or not an individual was recorded as working during the month following termination, during the third month following termination, and during the sixth month subsequent to termination. Similarly, the welfare status measures were calculated on the basis of activity during the last month of the time interval, and reflect whether or not an individual was recorded as receiving public assistance. Finally, the earnings measures are expressed on a weekly basis for both the three and six month periods following program termination, while the welfare benefit measures were estimated, in total, for the same time intervals.

4.2 Estimation of the Validation Benchmarks

As discussed earlier, the two validation benchmark measures used in this study are constructed from data on the participants in the AFDC Homemaker-Home Health Aide Demonstrations.[1] These measures include:

- The average monthly net increase in participant nonprogram earnings over the 32 months following program entry; and
- The average monthly net reduction in combined AFDC and food stamp benefits over the 29 months following program entry.[2]

In this section of the chapter, we describe the derivation of these measures, which occurred in two phases:

Phase 1: Estimation of overall program impacts and variations in impacts across subsets of participants (from the original demonstration study); and

Phase 2: Construction of individual-specific impact estimates (using the demonstration study results).

4.2.1 Impact Estimates From the Original Demonstration Study

As noted earlier, the AFDC Homemaker-Home Health Aide Demonstrations involved the random assignment of eligible applicants to either a treatment or control group. The 50-50 odds of assignment resulted in a total of 4,762 treatment group members and 4,758 control group members spread across seven states, including Arkansas, Kentucky, New Jersey, New York, Ohio, South Carolina, and Texas.

Program impacts were estimated in each state by comparing earnings and public benefit outcome levels in that state's treatment group to the same measures for the corresponding control group (which represented what would have happened to the treatment group absent the program). Two separate treatment/control comparisons were

---

[1] See Bell et al. (1987) for an overview of the demonstrations and related research findings.

[2] The follow-up data on AFDC and food stamp benefits were taken from a different source than the earnings data and limited to three fewer months. The follow-up period for participants in New York (six percent of the sample) was even more brief: 25 months for earnings and 26 months for AFDC and food stamp benefits. See Section 4.2.1 for more details on this.

made in each state, one for earnings outcomes taken from a telephone follow-up survey of 1,838 total treatment group members and 1,678 controls, and the other for public benefits, taken from welfare administrative records and covering 4,559 total treatment group members and 4,558 controls.

Data availability limited and defined each of these analysis samples: a small percentage of the treatment and control group members assigned could not be identified in state AFDC and food stamp program files and hence could not be included in the public benefit analysis; a much larger number of individuals were omitted from the earnings analysis either because they entered the program too late to be interviewed (4,162 cases) or because they could not be reached by survey staff (1,842 cases). Fortunately, exclusions from the earnings analysis sample appear to have occurred largely at random and, therefore, should not seriously bias the treatment group/control group comparison.[1]

Program impact estimates were calculated from these data using a regression model that adjusted the basic treatment/control comparison to take account of variations in impacts across different groups of participants. In its simplest form, the model can be stated as:

(4.2.1) $Y = \beta_0 + X\beta_1 + T(\delta_0 + X\delta_1) + \epsilon$, where

Y = the outcomes measure (nonprogram earnings or AFDC plus food stamp benefits);

X = a background variable that distinguishes one group of participants from another (e.g. a dummy variable that equals 1 for high school graduates and 0 for nongraduates);

T = a dummy variable that distinguishes members of the treatment group (T = 1) from members of the control group (T = 0); and

ϵ = a random error term.

Equation 4.2.1 predicts outcomes in the control group ($\beta_0 + \beta_1 X$) as a function of the baseline characteristics (X), and program impact ($\delta_0 + \delta_1 X$) as a function of the same baseline characteristics.

A large number of baseline variables were included in the original regression

---

[1]See Enns et al. (1987) for an indepth discussion of possible survey response bias in the nonprogram earnings analysis.

equations, both as separate variables (with β coefficients) and interacted with the treatment dummy (with δ coefficients).[1] Those variables are listed in Exhibit 4.1.[2] The inclusion of background variables in the impact model enables us to measure the δ_1 coefficient (or coefficients). This capability is of vital significance when estimating impacts for individual participants (e.g. the validation benchmarks), as described in the next subsection.

The impact regression in Equation 4.2.1 also included a set of dummy variables representing the various months following program entry, with a separate dummy variable for each month. When interacted with the treatment dummy, these variables allow us to estimate a different impact amount in each month after assignment, commensurate with monthly variations in outcome differences between the treatment and control groups.

Fourteen versions of Equation 4.2.1 were estimated, two for each state. The first equation in a state estimated effects on nonprogram earnings from the monthly earnings data, while the second equation estimated effects on combined AFDC and food stamp benefits from the monthly public benefits data. The sample size and average monthly impact for each regression is shown in Exhibit 4.2.[3]

4.2.2 Construction of Validation Benchmark Measures

For this study, we used the fourteen estimated regression equations from the original demonstration analysis to calculate the impact of the programs on each of the 1,548 participants for whom complete outcome data were available. Only actual training entrants are included in the analysis; individuals who were assigned to the treatment

---

[1]Enns et al. (1987) and Bell (1987) describe the original regression equations and estimation techniques in detail, for the earnings and public benefit analyses, respectively.

[2]Note that all of the variables in Exhibit 4.1 are determined either prior to random assignment (e.g. number of children in the AFDC case) or by factors external to the demonstration programs (e.g. local unemployment rate). Hence, they do not correlate with the treatment dummy, T, and cannot bias the basic treatment/control comparison.

[3]Inprogram earnings were also increased by the demonstrations. The combined effect of nonprogram earnings gains and inprogram earnings gains produced the public benefit reductions shown in Exhibit 4.2.2. Hence, public benefit reductions are larger than the nonprogram earnings gains alone.

EXHIBIT 4.1

BASELINE VARIABLES USED IN INITIAL IMPACT REGRESSIONS

1. Demographic Characteristics

   Age
   Race (white, black, Hispanic, Asian/Pacific, American Indian, other)
   Education (8th grade, 9th-11th grade, high school graduate, voc/tech training, some college, bachelor's degree)
   Marital status (married, widowed, separated/divorced, single)
   Number of children in AFDC case

2. Caregiving Experience

   Types of persons cared for other than own children (adults, children, both, neither)
   Number of persons cared for other than own children
   Paid for caregiving (yes, no)

3. Employment Experience

   Ever worked for pay (yes, no)
   Wage rate on last job (dollars per hour)
   Length of longest job (months)
   Time since last job (months)

4. Government Program Status

   WIN status (mandatory registrant, voluntary registrant, nonregistrant)
   Housing subsidy (yes, no)

5. Availability of Private Transportation

   Licensed driver (yes, no)
   Regular access to vehicle (yes, no)

6. Assessment of Potential as a Homemaker-Home Health Aide

   Education potential (1 = poor, 2 = fair, 3 = good, 4 = excellent)
   Work experience potential (1-4)
   Job-readiness potential (1-4)
   Personal potential (1-4)
   Overall potential (1-4)

7. Public Benefit History

   Average monthly AFDC payment over previous year (dollars)
   Change in AFDC payment in previous month (dollars)
   Average monthly food stamp payment over previous year (dollars)
   Change in food stamp payment in previous month (dollars)
   Time between first receipt of AFDC as an adult and program entry (months)

8. Programmatic/Environmental Factors

   Demonstration site (dummy variable for each site)
   Length of site operations prior to program entry (months)
   Local unemployment rate (percent)
   Average wage rate in local retail and service trades (dollars per hour)

EXHIBIT 4.2

SAMPLE SIZES AND AVERAGE MONTHLY IMPACTS FOR THE AFDC HOMEMAKER-HEALTH AIDE DEMONSTRATIONS, BY STATE AND OUTCOME MEASURE

State	Sample Size (months of data)		Monthly Impacts (dollars)	
	Earnings	Public Benefits	Earnings	Public Benefits
Arkansas	6,406	13,956	32***	-43***
Kentucky	7,485	14,517	13	-69***
New Jersey	18,978	40,228	42***	-80***
New York	3,774	8,930	-56***	3
Ohio	16,925	9,269[a]	20*	-92***
South Carolina	11,031	17,359	-15*	-101***
Texas	16,208	36,449	32***	-19**

*** Statistically significant at the 1 percent level.
** Statistically significant at the 5 percent level.
* Statistically signficiant at the 10 percent level.

[a] Public benefit effects in Ohio were estimated using data from only two counties (Butler and Hamilton). Food stamp data were not available for the rest of the state.

Sources: Enns et al. (1987), Tables III.4 and B.5; Bell (1987), Tables III.3 and B.5

group but who never began training are excluded. Two-hundred ninety of the 1,838 treatment group members with complete data fall into this group. The remaining participants -- referred to here as the trainee sample -- are distributed among states as indicated in Exhibit 4.3.

In focusing on training entrants, we first adjust the estimated coefficients in Equation 4.2.1 to take account of the "no-shows" -- those members of the treatment group who did not show up for training. Presumably, these individuals experienced no program effects. They were, however, included in the initial regression analyses, since their counterparts in the control group could not be precisely identified and excluded. As a result, the δ coefficients in the impact portion of Equation 4.2.1 are too small, reflecting a mixture of positive impacts for the trainees and zero impacts for the no-shows. We compensate for this bias by multiplying the δ coefficients in each state by the ratio of the number of individuals assigned to treatment to the number of training entrants, a ratio that always exceeds one.[1]

Once the coefficients are adjusted, we derive month-by-month impact estimates for each trainee using the formula:

(4.2.2) $I_{it} = \delta_{0t} + \delta_1 X_i$, where

I_{it} = program impact for trainee i in month t;

δ_{0t} = the adjusted intercept for month t in the impact component of Equation 4.2.1;[2]

δ_1 = the adjusted slope coefficient in the impact component of Equation 4.2.1; and

X_i = background characteristic for trainee i.

The month-by-month impact estimates are then averaged over time for each trainee to derive an individual's overall impact estimate, referred to here as that individual's validation benchmark:

(4.2.3) $$B_i = \frac{1}{M} \sum_{t=0}^{M} I_{it},$$

---

[1]Bloom (1984) showed that this adjustment exactly offsets the bias created by the inclusion of no-shows in the original regression analyses.

[2]The intercept term in Equation 4.2.2 varies with t through the time period dummy variables described above.

EXHIBIT 4.3

TRAINEES IN THE ANALYSIS SAMPLE, BY STATE

State	Number	Percent
Arkansas	123	8
Kentucky	155	10
New Jersey	364	24
New York	92	6
Ohio	326	21
South Carolina	220	14
Texas	268	17
All States	1,548	100

where t = 0 represents the month of program entry and t = M represents the last month with outcome data.

As noted earlier, separate earnings and public benefit benchmarks are calculated for each individual. The means and standard deviations of these two validation benchmark measures are shown in Exhibit 4.4, for all seven states combined. As can be seen, the AFDC Homemaker-Home Health Aide Demonstrations increased trainee nonprogram earnings about $18 per month, while reducing trainee public benefit payments about $66 per month.[1] More importantly, Exhibit 4.4 shows that both benchmark measures varied substantially within the sample of trainees, as indicated by their large standard deviations. This high level of variation will allow for more informative examinations of the relationship between the size of program impacts and the post-program outcome measures used in our validation analysis.

An important characteristic of the validation benchmarks is the extent to which they are correlated with each other. Intuitively, one would think that net increases in earnings should lead to net decreases in dependency on public welfare. This intuition underlies a major goal of JTPA as well as the use of employment and earnings related performance measures.

This did not turn out to be the case with the validation benchmarks estimated from AFDC Homemaker-Home Health Aide Demonstrations. While the correlation between the two benchmarks did have the correct sign, it was significant at only the 53 percent level. On the surface, this would imply that increases in earned income alone, may not be sufficient for inducing net reductions in welfare dependency. But why is this the case and how can we explain the lack of a strong and significant relationship between two outcomes otherwise believed to be strongly related? There are several possible explanations.

The first and most important explanation has to do with the construction of the welfare and earnings impact estimates themselves. During participation in the program, trainees received subsidized employment and wages. Because these earnings reduced welfare benefit levels, their effect was factored into the program's estimated impact on welfare dependency. However, because of their subsidized nature, program earnings

---

[1] As noted earlier, public benefit reductions exceed nonprogram earnings gains due to large (unreported) inprogram earnings gains.

EXHIBIT 4.4

MEANS AND STANDARD DEVIATIONS OF VALIDATION BENCHMARKS (in dollars)

	Program Impact on:	
	Monthly Nonprogram Earnings	Monthly Public Benefits
Mean	17.11	-66.32
Standard Deviation	139.42	167.40

Sample Size = 1,548

were not included in the estimates of earnings impacts. As a result, at least during the period of program participation, the earnings induced decline in welfare benefit levels was not matched with a measured increase in earned income. But since program participation accounted for only a small proportion of the total measurement period, this measurement issue would not necessarily be expected to eliminate the anticipated correlation completely. Other factors may have also been at work.

A second explanation is that it may be easier to favorably affect one's earnings than dependency on public assistance, especially since the majority of AFDC recipients ordinarily mix work and welfare. For example, most evaluations of the earnings impact of employment and training programs show positive net effects achieved largely through increases in employment, not wages.[1] These impacts, however, tend to be relatively small and not sufficient in size to make a large difference in one's welfare grant. Moreover, the small size of the impacts may also be insufficient to induce any change in the labor supply decisions of AFDC recipients, especially when compared to the size of the benefit level and that amount lost due to earnings disregards.

In the present case, the largest average monthly earnings impact was estimated to be $42 (Exhibit 4.2); on a cumulative annual basis, this translates into roughly $500. Given an average monthly payment level for a family of two of $382, the additional $42 cannot have a large effect on one's grant level; nor would it appear to be sufficient in size to induce a large change in one's willingness to work.[2] As a result, one would expect a limited correlation between earnings gains and welfare reductions.

The third explanation is that earnings impacts may be much less sensitive to targeting decisions than are reductions in welfare dependency. It could well be that even for the relatively less disadvantaged, it is more possible to produce a modest earnings impact than it is to produce an impact on welfare dependency. Indeed, as we discuss in Chapter 5.0, the production of favorable welfare impacts is much more sensitive to one's prior work and welfare history than is achieving gains in one's earnings. This suggests that the simultaneous achievement of both objectives may be dependent on the background characteristics of those served.

---

[1]This was also found in the evaluation of the AFDC Home-Health-Aide Demonstration. See: Enns et al. (1987) for a review of the results.

[2]Based on a family of one adult recipient and two children between October 1985 and September 1986.

A fourth and final explanation is that welfare impacts have an upper limit, defined at the point when one completely moves off the public assistance rolls. In contrast, earnings impacts are, theoretically, unlimited. As a result, these two facts will tend to limit the overall relationship between earnings and welfare impacts, although they should not eliminate the correlation altogether.

All of this is not intended to imply that the pursuit of earnings gains is inconsistent with the goal of reducing welfare dependency. It may simply mean that improvements in one's earnings status are not sufficient for producing net reductions in welfare dependency. As discussed above, other factors, including the choice of whom to serve, can also be expected to play an important role.

4.3 Validation Tests Conducted

To assess the validity of the candidate performance measures in predicting each of the two benchmarks, we tested them alone and with respect to their incremental addition to the predictive validity of others. To test the measures alone, we simply correlated each one with the two benchmarks. For example, to determine the validity of employment status at six months after program termination, we correlated each of the two benchmarks with a variable indicating whether or not a trainee was employed during the sixth month following termination. The resulting correlation coefficients and their accompanying levels of statistical significance provide an indication of the relative predictive validity of each measure when used alone.

In practice, however, more than one performance measure may be used at the same time. For this reason, we also assessed the additional or incremental predictive validity of a measure when added to another. This analysis was conducted to address two key questions: 1) by how much do post-program measures add to the validity of employment status at the time of termination? and 2) by how much does increasing the post-program period following termination increase predictive value?

In addition to these independent and incremental tests, we also examined the correlation between the various indicators themselves. This was done to determine how complementary the candidate measures are with each other, and to identify the extent to which it may be appropriate to rely on some combination of welfare, employment, or earnings performance measures.

Prior to actually conducting the validation tests, one further step was judged necessary. This step adjusted each candidate performance measure for differences in

individual background characteristics. The adjustment procedure was followed to reflect the manner by which performance standards are used in the JTPA program and recommended for use in the welfare reform proposals.

Currently, the Department of Labor offers states the option of adjusting local SDA performance standards to reflect differences in client mix and economic conditions. This adjustment is achieved through the use of a multiple regression model estimated from data on the performance, client characteristics, and economic conditions in each SDA. In this way, SDAs are held-harmless from factors which influence performance but over which they have little if any control.

To mirror this procedure, we estimated regression models for each of the ten candidate performance measures included in the validation tests.[1] Using each candidate measure as a dependent variable, the following factors were entered into the regression equation as independent variables: 1) age; 2) education; 3) race; 4) marital status; 5) number of children in the AFDC case; 6) number of continuous months of unemployment prior to program application; 7) pre-program welfare dependency; 8) state unemployment rate; and 9) state average hourly wage rate.

Having estimated the adjustment models, the final step was to construct the test variables. This was accomplished by calculating the residual from each regression model for each individual trainee in the analysis. For example, the test variable for average weekly earnings was calculated as follows. First, this measure was regressed on the nine independent variables listed above. Then, the difference between the actual observed value of the measure and that predicted by the regression was calculated to compute the residual. This procedure was repeated for all of the candidate measures, purging from the validation tests any correlation that might exist between the independent variables, the candidate measures, and the validation benchmarks.

Having accomplished this task, we then proceeded to conduct the validation tests. Our results are presented in the next section.

4.4 Key Validation Findings

When examining and attempting to interpret the validation results presented in this section, it is important to bear in mind two key points. The first is that the absolute

---

[1] The regression results are reported in Appendix 4-A.

magnitude of the validation correlations is less important than their relative values because the objective of the tests is to make a selection from among a specified set of alternatives. Thus, the appropriate focus should be on the pattern of the validation correlations as opposed to their specific, absolute values.

Further, one should also expect the absolute value of the correlations to be relatively small. Because we are using individual micro-based data, and since we are attempting to explain variations in individual impacts with only one or a few potentially important explanatory variables, it is not realistic to expect them to account for a major share of the variation in the benchmarks.

The second point has to do with clarifying the expected relationship between the welfare reductions benchmark and the candidate indicators. As presented, the welfare reduction benchmark measures the net decrease in public welfare benefits received by the trainees as a result of their participation in the program. Since the disregard provisions of the AFDC program require a reduction in benefits levels once a recipient goes to work, one would expect the employment and earnings indicators to be negatively correlated with the welfare dependency benchmark. Additionally, this correlation should increase with the post-program period since the earnings disregards are more stringent after as opposed to during the first three months of employment. As currently designed, welfare benefits are taxed at a rate of 33 percent during the first three months that a recipient goes to work. After that time, the tax rate on benefits as a result of working increases to roughly 100 percent; in other words, for each dollar earned, there is a dollar reduction in benefit levels.

With this in mind, Exhibit 4.5 summarizes the findings obtained by comparing the ten candidate performance measures with each of the two validation benchmarks. As can be seen, the correlation coefficients tend to be relatively small, ranging in absolute value from .005 to .151. In large part, this should be expected given the many other factors, besides performance outcomes, that influence earnings gains and welfare reductions. What is notable, however, is the large number of coefficients that have the correct sign and are statistically significant.

Overall, of the twenty separate correlations tested, fourteen or 70 percent are significant with the expected sign, suggesting that there are viable candidate performance measures from which to choose. However, the findings also indicate that earnings gains are predicted by a wider range of potential performance measures than net reductions in welfare dependency. While all of the employment and earnings indicators are valid predictors of earnings gains, only two are significantly related to the

EXHIBIT 4.5

CORRELATION RESULTS USING

ADJUSTED PERFORMANCE MEASURES[a]

CANDIDATE MEASURES	VALIDATION BENCHMARK: EARNINGS GAINS	VALIDATION BENCHMARK: REDUCTIONS IN WELFARE DEPENDENCY
1. Placed at Termination (1=placed; 0=otherwise)	0.051 (0.053)	-0.006 (0.813)
2. Average Placement Wage	0.083 (0.002)	-0.020 (0.438)
3. Employment Status (1=employment; 0=otherwise)		
3 months	0.079 (0.003)	-0.033 (0.202)
6 months	0.126 (0.000)	-0.042 (0.108)
4. Average Weekly Earnings		
3 months	0.120 (0.000)	-0.032 (0.222)
6 months	0.151 (0.000)	-0.046 (0.080)
5. Welfare Status (1=on welfare; 0=otherwise)		
3 months	-0.005 (0.839)	0.063 (0.017)
6 months	-0.017 (0.517)	0.089 (0.001)
6. Welfare Benefits Level		
3 months	-0.043 (0.098)	0.090 (0.001)
6 months	-0.047 (0.074)	0.101 (0.000)

[a] Levels of significance are reported in parenthesis

welfare benchmark. As would be expected, those measures reflecting one's welfare status and benefit level best predict net reductions in public assistance.

These findings largely reflect the limited correlation between the two validation benchmarks. Fortunately, it is possible to identify certain measures that are valid for both outcomes of interest. In particular, the findings in Exhibit 4.5 show that average weekly earnings over the six months following program termination, and welfare status and benefit levels at three and six months are significantly correlated with each benchmark measure. In contrast to the popular preference for and intuitive appeal of one's placement status following termination, this performance measure does not appear to be a valid predictor of reductions in welfare dependency.

The relatively limited validity of placement status is also evident when compared to the post-program measures included in the analysis. As the results show, the predictive validity of the candidate measures increases as the post-program period is lengthened. For both benchmarks, post-program measures out-perform the two termination-based measures; and of the post-program indicators, those measured over six months do better than those calculated at the three month termination point. By and large, the best predictor of earnings gains is average weekly earnings over the six month period following termination. While this measure is also a strong predictor of welfare reductions, it does appear to be slightly less predictive than either of the two welfare-based indicators.

This is not to say that the three months measures are invalid. To the contrary, employment status, average weekly earnings, and welfare benefit levels at three months are each significantly correlated with the earnings benchmark. While the magnitude of the correlation is smaller than that at six months, the difference tends to be relatively small. The same is true about the correlations between the two welfare indicators and the welfare reductions benchmark; while six months out performs the three month measures, the difference is not large.

The only exception is the correlation between the employment and earnings indicators and the welfare benchmark. Here, the three month measures are not only a bit smaller than the six month indicators, but they are also statistically insignificant. This is likely the result of the fact that the implicit AFDC tax rate on earnings increases after the first three months of earnings.

Since these results are based solely on comparative judgments, we reassessed them using a statistical procedure designed to identify that candidate measure which

yielded the greatest predictive power, commonly referred to as the R^2. This procedure was employed to identify the best indicator from those measuring employment status, earnings, welfare status, welfare benefit levels, respectively. As a final test, the procedure was relied on to identify the best indicator from among all of those tested. The test was run first for the earnings gain benchmark and then for the welfare reduction benchmark.

The results of these tests are presented in Exhibit 4.6 and show a high degree of similarity with those just reported. In addition, the results also show a strong consistency between those indicators in each class of performance measures that best predict the two validation benchmarks. But perhaps most important is the consistent selection of six month post-program measures as the best single predictor of both earnings gains and welfare reductions. Regardless of the type of indicator (e.g. employment status, welfare benefit level), or the benchmark, six month performance measures are better predictors than other indicators measured at either the three month point or at the time of termination. Also important is the selection of average weekly earnings and average monthly benefit levels as the overall best predictors of earnings gains and welfare reductions, respectively.

Although placement status did not turn out to be a valid predictor of welfare reductions, it did significantly correlate with earnings gains. Because of this, we also evaluated the incremental validity of using other indicators in combination with one's placement status at the time of termination. More specifically, we employed the same procedure as above, but in this case, took one's placement status as given, thus allowing us to readily determine the incremental validity of adding an additional candidate measure.

The results of this test are presented in Exhibit 4.7. First and for the earnings gain benchmark, there is little support for using the placement status indicator. As can be seen, while the predictive validity of two indicators is superior to using just one, little additional information is gained from the combination; this is evident from the fact that the coefficient for the placement indicators changes sign and becomes insignificant in the presence of the other measures. We do, however, obtain the correct sign of the placement indicator when used in combination with a welfare-related performance measure, although it remains insignificant.

Taken together, these results have several important implications for the selection of performance measures for programs targeted on serving AFDC recipients. The implications can be summarized as follows:

EXHIBIT 4.6

ADJUSTED CANDIDATE MEASURES WITH GREATEST PREDICTIVE VALIDITY, BY BENCHMARK

	VALIDATION BENCHMARK									
	EARNINGS GAINS					REDUCTIONS IN WELFARE DEPENDENCY				
	Best Emp. Indicator	Best Earn. Indicator	Best E&E Indicator	Best Welf. Indicator	Best of All	Best Emp. Indicator	Best Earn. Indicator	Best E&E Indicator	Best Welf. Indicator	Best of All
Employment Indicators										
Placement Status										
Employment Status										
3 months										
6 months	40.489[a]					-16.854				
Earnings Indicators										
Average Placement Wage										
Average Weekly Earnings										
3 months										
6 months		0.271[a]	0.271[a]		0.271[a]		-0.103[c]	-0.103[c]		
Welfare Indicators										
Welfare Status										
3 months										
6 months										
Welfare Benefits Level										
3 months										
6 months				-0.008[c]					0.020[a]	0.020[a]
Multiple Correlation Coefficient ($\sqrt{R^2}$)	0.126[a]	0.151[a]	0.151[a]	.047[c]	0.151[a]	0.042	0.046[c]	0.046[c]	0.101[a]	0.101[a]

Note: Indicator entries reflect parameter estimates from regressions equations

Level of Significance

[a] Significant at 99%
[b] Significant at 95%
[c] Significant at 90%

EXHIBIT 4.7

ADJUSTED CANDIDATE MEASURES WITH GREATEST PREDICTIVE VALIDITY: PLACEMENT STATUS GIVEN

	VALIDATION BENCHMARK: EARNINGS GAINS						VALIDATION BENCHMARK: REDUCTIONS IN WELFARE DEPENDENCY					
	Placement Status Alone	Best Emp. Indicator	Best Earn. Indicator	Best E&E Indicator	Best Welf. Indicator	Best of All	Placement Status Alone	Best Emp. Indicator	Best Earn. Indicator	Best E&E Indicator	Best Welf. Indicator	Best of All
Employment Indicators												
Placement Status	-	-3.488	-22.852[c]	-22.852[c]	15.726	-22.852[c]	-	8.078	15.526	15.526	13.378	13.378
Employment Status												
3 months												
6 months		41.795[a]						-19.878[c]				
Earnings Indicators												
Average Placement Wage												
Average Weekly Earnings												
3 months												
6 months			0.332[a]	0.332[a]		0.332[a]			-0.144[b]	-0.144[b]		
Welfare Indicators												
Welfare Status												
3 months												
6 months												
Welfare Benefits Level												
3 months												
6 months					-0.006						0.022[a]	0.022[a]
Multiple Correlation Coefficient $\sqrt{R^2}$	.051[b]	0.126[a]	0.158[a]	0.158[a]	0.060[c]	0.158[a]	.006	0.045	0.053	0.053	0.104[a]	0.104[a]

Note: Indicator entries reflect parameter estimates from regressions equations

Level of Significance

[a] Significant at 99%
[b] Significant at 95%
[c] Significant at 90%

1. The vast majority of candidate measures tested do have potential value as performance indicators. Of the twenty indicators tested, fourteen or 70 percent are found to be statistically significant with the correct sign.

2. Earnings gains are considerably easier to predict than reductions in welfare dependency. While all but two candidate measures were correlated with the earnings benchmark, five of the candidate measures were uncorrelated with the welfare reductions benchmark.

3. Post-program measures consistently have greater predictive validity than either of the two termination-based measures. However, while the six month indicators out perform those measured at three months, the difference in magnitude is often small.

4. Of all the measures tested, average weekly earnings over the six months following termination and average monthly benefit levels, also over six months, best predict both validation benchmarks.

These findings point to the potential desirability of using a combination of earnings and welfare benefit levels as performance measures for programs targeted on serving AFDC recipients. Two issues remain, however, for consideration. The first is empirical in nature and focused on the degree of complementarity that exists between the two types of measures. If earnings and welfare benefit levels both appear to be valid predictors of net reductions in welfare dependency and increases in earnings, then it is important to demonstrate that the two are compatible with each other.

To assess this, we examined the correlation between the candidate indicators included in the analysis. The results are presented in Exhibit 4.8 and reveal several encouraging results. First, the findings show why the use of more than one employment and earnings indicator yields little additional information than that obtained from one indicator alone. As can be seen, there is a very high degree of correlation between and among the six employment and earnings candidate measures tested. In some cases, the correlation coefficient exceeds .80, while in no case is it less than .46.

Second, while the results are similar for the welfare-based indicators, the level of correlation is not as great. Moreover and importantly, there does appear to be a significant correlation between the employment and earnings indicators and the welfare indicators. This finding holds for all of the indicators tested, as well as for the relationship between six month earnings and six month welfare status or welfare benefit level. Thus, we can say that the use of these two indicators will be mutually reinforcing. That is, promoting the post-program earnings of program trainees should

EXHIBIT 4.8

DEGREE OF CORRELATION AMONG ADJUSTED CANDIDATE MEASURES[a]

			Employment Status		Earnings Status		Welfare Status		Welfare Benefits Level	
	Placed at Termination	Average Placement Wage	3 months	6 months	3 months	6 months	3 months	6 months	3 months	6 months
Placed at Termination	1.000 (0.000)	0.841 (0.0001)	0.615 (0.0001)	0.458 (0.0001)	0.633 (0.0001)	0.588 (0.0001)	-0.254 (0.0001)	-0.158 (0.0001)	-0.298 (0.0001)	-0.307 (0.0001)
Average Placement Wage	0.841 (0.0001)	1.000 (0.0000)	0.637 (0.0001)	0.481 (0.0001)	0.758 (0.0001)	0.712 (0.0001)	-0.289 (0.0001)	-0.175 (0.0001)	-0.341 (0.0001)	-0.347 (0.0001)
Employment Status										
3 months	0.615 (0.0001)	0.637 (0.0001)	1.000 (0.0000)	0.713 (0.0001)	0.805 (0.0001)	0.804 (0.0001)	-0.296 (0.0001)	-0.271 (0.0001)	-0.354 (0.0001)	-0.415 (0.0001)
6 months	0.458 (0.0001)	0.481 (0.0001)	0.713 (0.0001)	1.000 (0.0000)	0.591 (0.0001)	0.731 (0.0001)	-0.247 (0.0001)	-0.257 (0.0001)	-0.316 (0.0001)	-0.368 (0.0001)
Earnings Status										
3 months	0.633 (0.0001)	0.758 (0.0001)	0.805 (0.0001)	0.591 (0.0001)	1.000 (0.0000)	0.955 (0.0001)	-0.326 (0.0001)	-0.261 (0.0001)	-0.374 (0.0001)	-0.421 (0.0001)
6 months	0.588 (0.0001)	0.712 (0.0001)	0.804 (0.0001)	0.731 (0.0001)	0.955 (0.0001)	1.000 (0.0000)	-0.321 (0.0001)	-0.293 (0.0001)	-0.375 (0.0001)	-0.437 (0.0001)
Welfare Status										
3 months	-0.254 (0.0001)	-0.289 (0.0001)	-0.296 (0.0001)	-0.247 (0.0001)	-0.326 (0.0001)	-0.321 (0.0001)	1.000 (0.0000)	0.542 (0.0001)	0.618 (0.0001)	0.612 (0.0001)
6 months	-0.158 (0.0001)	-0.175 (0.0001)	-0.271 (0.0001)	-0.257 (0.0001)	-0.261 (0.0001)	-0.293 (0.0001)	0.542 (0.0001)	1.000 (0.0000)	0.374 (0.0001)	0.520 (0.0001)
Welfare Benefits Level										
3 months	-0.298 (0.0001)	-0.341 (0.0001)	-0.354 (0.0001)	-0.316 (0.0001)	-0.374 (0.0001)	-0.375 (0.0001)	0.618 (0.0001)	0.374 (0.0001)	1.000 (0.0000)	0.939 (0.0001)
6 months	-0.307 (0.0001)	-0.347 (0.0001)	-0.415 (0.0001)	-0.368 (0.0001)	-0.421 (0.0001)	-0.437 (0.0001)	0.612 (0.0001)	0.520 (0.0001)	0.939 (0.0001)	1.000 (0.0000)

[a] Levels of significance are reported in parenthesis

lead to favorable short-term welfare results and longer-term gains in earnings in reductions in welfare dependency.

But the most statistically valid indicator may not be the most practical one to measure and monitor. The feasibility and cost of data collection must play an important role in making the final selection of candidate indicators. Although the six month indicators seem to perform the best, they are more difficult to measure and require longer lead time to use for management purposes. As was seen, the three month measures predicted earnings gains nearly as well as the six month measures. Similarly, while monthly benefit levels outperform measure of welfare status, the former is more difficult to obtain. Thus, practical considerations may suggest a preference to welfare status measures.

Given this, the tradeoffs are fairly clear. If the sole objective is to select those measures that are the best predictors of both benchmarks, then earnings and welfare benefit levels over six months would be most appropriate. However, given practical considerations, other choices are also available. For example, earnings over the three months following program termination is nearly just as good a predictor of the earnings benchmark as when measured at six months. Although not a valid predictor of reductions in welfare dependency, the three month earnings measure is strongly correlated with the candidate welfare measures, providing the needed degree of inter-indicators consistency.

Overall, the results presented in this chapter provide a range of valid indicators from which to choose. Ultimately, tradeoffs between the feasibility of data collection and management utility on the one hand, and predictive validity on the other hand, will have to guide the final selection of performance measures for targeted welfare programs.

APPENDIX 4-A

EXHIBIT 4-A.1

PERFORMANCE INDICATOR ADJUSTMENT MODELS

	Employment Measures			Earnings Measures			Welfare Measures			
	Placement	3 mos	6 mos	Placement	3 mos	6 mos	Status 3 mos	Status 6 mos	Benefits 3 mos	Benefits 6 mos
Intercept	0.328[a]	0.590[a]	0.561[a]	1.863[a]	94.578[a]	96.458[a]	0.503[a]	0.604[a]	206.709[a]	506.417[a]
	(6.203)	(8.271)	(8.680)	(5.350)	(7.512)	(8.337)	(7.366)	(9.134)	(3.026)	(127.617)
Age	0.001	0.000	0.001	0.021[b]	0.296	0.281	-0.001	-0.001	-4.261[a]	-9.129[a]
	(0.591)	(0.032)	(0.306)	(2.336)	(0.928)	(0.959)	(-0.728)	(-0.401)	(-2.639)	(-3.028)
< 9 yrs. educ.	-0.037	-0.086[c]	-0.079[c]	-0.538[b]	-18.360[b]	-18.687[b]	0.066	-0.001	2.914	11.932
	(-1.074)	(-1.833)	(-1.863)	(-2.363)	(-2.230)	(-2.470)	(1.471)	(-0.013)	(0.065)	(0.142)
9-11 yrs. educ.	-0.043[c]	-0.045	-0.036	-0.403[a]	-12.067[b]	-10.335[b]	-0.018	-0.004	-8.216	-8.618
	(-1.944)	(-1.506)	(-1.322)	(-2.769)	(-2.293)	(-2.137)	(-0.619)	(-0.152)	(-0.286)	(-0.160)
> 12 yrs educ.	0.001	0.031	0.054[c]	0.061	9.902[c]	9.900[c]	-0.042	0.004	-57.808[c]	-102.094[c]
	(0.042)	(0.927)	(1.745)	(0.370)	(1.658)	(1.804)	(-1.287)	(0.134)	(-1.766)	(-1.670)
Black	-0.069[a]	-0.101[a]	-0.085[a]	-0.547[a]	-22.214[a]	-20.332[a]	0.060[b]	0.068[b]	34.565	88.325
	(-2.891)	(-3.137)	(-2.904)	(-3.481)	(-3.912)	(-3.896)	(1.947)	(2.280)	(1.114)	(1.524)
Other	-0.015	-0.034	-0.084[b]	-0.236	-11.870	-15.175[b]	0.018	-0.022	64.556	105.464
	(-0.438)	(-0.752)	(-2.047)	(-1.070)	(-1.486)	(-2.068)	(0.415)	-0.515	(1.473)	(1.288)
Married	0.062[c]	-0.028	-0.008	0.284	-2.233	-2.782	-0.065	0.005	-77.862[c]	-161.362[b]
	(1.833)	(-0.624)	(-0.186)	(1.285)	(-0.279)	(-0.378)	(-1.483)	(0.107)	(-1.774)	(-1.968)
Single	-0.049[b]	-0.055[c]	-0.034	-0.326[b]	-14.165[a]	-11.977[b]	0.078[a]	0.068[b]	66.308[b]	126.110[b]
	(-2.177)	(-1.819)	(-1.254)	(-2.217)	(-2.667)	(-2.454)	(2.687)	(2.393)	(2.272)	(2.313)
# children in AFDC case	-0.011	-0.017	-0.017	-0.067	-3.278	-4.181[b]	0.053[a]	0.036[a]	154.020[a]	299.794[a]
	(-1.258)	(-1.456)	(-1.638)	(-1.196)	(-1.630)	(-2.262)	(4.074)	(2.863)	(11.900)	(12.400)
Pre-prog. mon. unemp.	0.000	-0.001[b]	-0.001[b]	-0.002	-0.167[a]	-0.163[a]	0.000	0.000	0.130	0.506
	(-0.680)	(-2.069)	(-2.271)	(-1.566)	(-2.915)	(-3.103)	(0.852)	(0.796)	(0.409)	(0.850)
Pre-prog. ave. monthly benefits							0.000[a]	0.000[c]	0.471[a]	0.829[a]
							(2.970)	(1.823)	(5.261)	(4.960)
Total months rec'd AFDC	0.000	0.000	0.000[c]	0.001	0.043	0.039				
	(0.732)	(1.434)	(1.833)	(0.646)	(1.083)	(1.053)				

EXHIBIT 4-A.1 (continued)

PERFORMANCE INDICATOR ADJUSTMENT MODELS

	Employment Measures			Earnings Measures			Welfare Measures			
	Placement	3 mos	6 mos	Placement	3 mos	6 mos	Status 3 mos	Status 6 mos	Benefits 3 mos	Benefits 6 mos
Site unempl. rate	-0.006[c]	-0.011[a]	-0.011[a]	-0.061[a]	-2.408[a]	-2.557[a]	0.012[a]	0.005	5.477	5.093
	(-1.842)	(-2.671)	(-2.948)	(-3.035)	(-3.338)	(-3.859)	(3.137)	(1.401)	(1.382)	(0.688)
Site ave. wage	0.000[b]	0.000[b]	0.000[a]	0.000[a]	0.013[a]	0.0132[a]	0.000[a]	0.000[a]	-0.046[b]	-0.074[b]
	(2.012)	(2.508)	(2.613)	(3.772)	(3.389)	(3.881)	(-4.452)	(-3.166)	(-2.266)	(-1.967)
R^2	0.026	0.026	0.028	0.043	0.047	0.0501	0.0601	0.0356	0.198	0.207
F	2.908[a]	2.923[a]	3.157[a]	5.016[a]	5.424[a]	5.855[a]	7.092[a]	4.094[a]	27.440[a]	28.969[a]

[a] Significant at 99%
[b] Significant at 95%
[c] Significant at 90%

5.0 ESTABLISHING A PERFORMANCE MANAGEMENT SYSTEM

The development of a flexible and informative performance management system that supports the delivery of services to the AFDC population needs to go beyond the identification of statistically valid indicators. The validity of the indicators themselves may not be insufficient for making a final selection since the possibility exists that they may introduce incentives to cream from among the eligible population and thereby offer relatively low cost, short-term services. To the extent that service providers -- SDAs and contractors alike -- are motivated to achieve their performance benchmarks, they have some incentive to select individuals and offer program services that maximize their chances of meeting their management objectives. Indeed, this has been a major criticism of the performance management approach used currently in JTPA and formerly in other employment and training programs. It is thus quite important to consider this possibility and develop those management tools necessary to limit its occurrence.

In the JTPA program, the optional adjustment model developed by the Department of Labor is the chief mechanism available to limit the effects that performance standards may have on targeting and program mix decisions. The rationale underlying the adjustment methodology is to remove from the standards-setting process those factors which affect performance but over which SDAs have little if any control. In particular, by accounting for client characteristics, it is expected that preferences to enroll less disadvantaged individuals and/or provide low cost programs will be substantially limited.

Whether or not this approach is effective has come under increasing scrutiny in the recent past. In addition to the concerns raised at the federal level, states have indicated that the available adjustment mechanisms have at best, a limited influence on the program and targeting preferences of SDAs and their service providers. As discussed in our interim report, the small size of the adjustment factors, the relative ease by which SDAs can meet their performance expectations, and the voluntary nature of using the adjustments contribute to limiting the neutralizing effects that were expected from the performance standards adjustment process.[1] It may well be that the performance

---

[1]SDA performance expectations have been established so that roughly 75 percent are expected to be able to achieve them. For a detailed discussion of this, see: Jeffrey Zornitsky and Mary Rubin, Serving AFDC Recipients: Initial Findings on the Role of Performance Standards, Abt Associates Inc., Cambridge, MA, April, 1988.

management system will have to be changed or modified to better support service delivery to the AFDC population.

In this chapter of the report, we address these issues by first considering selected properties of the candidate measures themselves. In particular, we begin by exploring the potential for selective enrollment by examining the relationship between background characteristics that may be used for "creaming" (e.g. prior work and welfare dependency history) and both the candidate measures and the two validation benchmarks. To the extent that creaming incentives exist, we would expect to find a clear and significant correlation between these background characteristics and the long-term outcomes of interest.

While establishing that such relationships exist will not be new information per se, it will clearly indicate the importance of targeting decisions to the achievement of both the performance measures and the longer-term impacts expected of the program. Although adjustment models may represent one tool for addressing this, it may be that more specific targeting criteria can better limit any perverse incentives introduced by the performance standards.

But regardless of the statistical relationships that may exist between performance achievement, targeting, and program mix decisions, it remains important to gain some understanding of their relative importance. A growing presumption underlying much of the policy discussion over how to improve performance standards is that "the right changes" will induce program administrators and their contractors to devote more attention to serving AFDC recipients, and particularly those deemed hard-to-employ. However, several other key variables, such as child care and health benefits as well as inter-agency coordination, bear upon ability of employment and training programs to target on such individuals and provide them with the mix of services needed to ensure gainful employment.[1] Unless efforts are also made to address these issues, modifications to the performance standards system may have only a limited effect on future targeting and programming decisions.

It is thus quite important also to consider how SDAs plan their programs and targeting decisions, and the role that performance standards play in this process. In this way, it is possible to gain a better understanding of the context in which performance

---

[1]See Chapter Two for a discussion of these variables.

standards are managed and those changes that may be required to better stimulate service delivery to the AFDC population. This issue is addressed in Section 5.2 of this chapter and relies on the findings obtained from both our telephone interviews with state officials and field visits to SDAs and state offices.

On the basis of the results obtained from these two tasks, we then turn our attention to a discussion of options for developing a performance management system. Specifically, we address the viability of the modeling approach, improvements in the incentive structure, and the use of targeting criteria in the performance standards setting process.

5.1 Performance Standards and Creaming: Further Findings from the Validation Tests

To assess the potential creaming incentives that performance standards may introduce, we adopted an approach that examined various aspects of the correlation between selected background characteristics, the candidate measures, and the two validation benchmarks. Overall, our findings do reveal strong and consistent relationships between variables that could be used for selective enrollment and both the candidate performance measures and the two validation benchmarks. This is especially evident for net reductions in welfare dependency, suggesting that creaming is indeed a viable vehicle to ensure performance success.

One way to assess this is by re-examining the validation results reported in the previous chapter using unadjusted performance measures. Recall, that in conducting the validation tests we adjusted the candidate measures to remove the influence of individual background characteristics. By leaving the effect of these characteristics in the candidate measures, we can indirectly observe the effect they have on the results. If the creaming argument is correct, we would expect the predictive power of the measures to decrease or even be insignificant, given the positive relationship between net program impacts and the degree of disadvantagedness.[1]

The results of this re-test are presented in Exhibit 5.1 and partly support the argument of potential creaming. If we examine the findings for the earnings benchmark

---

[1]This test may result in overstated findings since individual background characteristics (and other variables) were used to estimate the benchmarks (see Section 4.2). However, since the number of these characteristics far exceeds those assessed here, the basic findings should hold.

EXHIBIT 5.1

CORRELATION RESULTS USING UNADJUSTED AND

ADJUSTED PERFORMANCE MEASURES[a]

CANDIDATE MEASURES	VALIDATION BENCHMARK			
	EARNINGS GAINS		REDUCTIONS IN WELFARE DEPENDENCY	
	Unadjusted	Adjusted	Unadjusted	Adjusted
1. Placed at Termination	0.053 (0.036)	0.051 (0.053)	0.027 (0.295)	-0.006 (0.813)
2. Average Placement Wage	0.086 (0.001)	0.083 (0.002)	0.041 (0.109)	-0.020 (0.438)
3. Employment Status				
3 months	0.068 (0.007)	0.079 (0.003)	0.007 (0.783)	-0.033 (0.202)
6 months	0.119 (0.000)	0.126 (0.000)	-0.003 (0.920)	-0.042 (0.108)
4. Average Weekly Earnings				
3 months	0.113 (0.000)	0.120 (0.000)	0.023 (0.363)	-0.032 (0.222)
6 months	0.142 (0.000)	0.151 (0.000)	0.011 (0.661)	-0.046 (0.080)
5. Welfare Status				
3 months	-0.017 (0.506)	-0.005 (0.839)	-0.010 (0.687)	0.063 (0.017)
6 months	-0.024 (0.344)	-0.017 (0.517)	0.017 (0.511)	0.089 (0.001)
6. Welfare Benefits Level				
3 months	-0.046 (0.072)	-0.043 (0.098)	0.046 (0.073)	0.090 (0.001)
6 months	-0.048 (0.059)	-0.047 (0.074)	0.050 (0.048)	0.101 (0.000)

[a] Levels of significance reported in parenthesis

first, the exhibit shows that the predictive validity of the candidate measures is not sensitive to the presence or absence of background characteristics. The same eight adjusted indicators that were found to be significant remain valid when measured in an unadjusted fashion. This finding does not, however, mean that individual background characteristics are unrelated to the employment and earnings performance measures. Indeed, as we show later, key variables that affect one's employability also affect the performance measures of interest. What the result likely reflects is the fact that the employment and earnings candidate measures are, by definition, direct components of the earnings benchmark.

In contrast to these results, the findings for the welfare reductions benchmark do reveal substantial sensitivity to the presence of background characteristics in the candidate measures. Overall, we observe three types of improvements in the validation results when they are conducted with adjusted measures. First, the vast majority of unadjusted measures have the incorrect sign; once they are adjusted, the expected relationships assume the correct sign. For example, one would expect higher post-program earnings to lead to net reductions in welfare dependency. When this proposition is tested using the unadjusted measures, the results suggest just the opposite. However, once the earnings indicators are adjusted for background characteristics, we observe the expected result. As the exhibit shows, this type of improvement occurred for six of the ten measures tested.

Second, adjusting the candidate measures also improves the statistical significance of the indicators. Of the six adjusted indicators found to be significant, all reflected a higher degree of significance than their unadjusted counterparts. In fact, four turned from insignificant to significant as a result of the adjustment.

The third observed improvement includes increases in the value of the correlation coefficients. The results clearly show that the predictive power of the indicators increases once they are expressed in an adjusted fashion.

The sensitivity of the welfare reductions benchmark to the presence (or absence) of individual background characteristics has two important implications. First, as mentioned earlier, it implies the presence of creaming incentives when the indicators are unadjusted. However, even when the indicators are adjusted, it is unclear whether such incentives are truly removed. Recall that our discussion with state and SDA practitioners indicated the limited effect of the DOL model on program decisions due to the small size of the adjustment factors. Moreover, since the choice of who to serve is left to states and SDAs, indicator adjustments can only have an influence when a decision

is made by an SDA (or service provider) to target a hard-to-employ group. Thus, the presence of an adjustment mechanism alone, will not necessarily be sufficient to stimulate services to AFDC recipients or those identified as hard-to-employ.

The second implication has to do with the connection between targeting and the generation of net impacts. The results observed in Exhibit 5.1 at least suggest background characteristics are correlated with net impacts. Since most of the unadjusted indicators were not correlated with the welfare reductions benchmark, one can tentatively conclude that this was caused by the strong and inverse association between high performance scores and one's level of disadvantagedness. For example, while one's attachment to the labor force is positively associated with the likelihood of employment, it is negatively associated with the level of impacts on welfare dependency; that is, those with substantial pre-program work experience probably have a strong chance of finding and holding a job in the post-program period, but a relatively limited opportunity to experience a net reduction in welfare dependency for a given level of program investment. Thus, while selective enrollment may facilitate performance achievement, it may well run counter to producing net impacts. Net impacts appear best achieved when programs are targeted on relatively disadvantaged individuals.

These issues are more directly assessed in Exhibits 5.2 and 5.3. In the first exhibit, we evaluated the correlation between a number of pre-program work history and welfare dependency characteristics and the two validation benchmarks. As expected, these characteristics are correlated with the benchmarks.[1] If we consider earnings gains first, it is evident that each of the six background characteristics measured have the correct sign and are statistically significant. The same is generally true for the welfare reductions benchmark, although two of the six variables are not significant. Just as important, however, is the fact that the strength of the correlations between these variables and the two benchmarks is as great or greater than that observed for the candidate performance measures. It would thus appear that targeting criteria or selective enrollment can indeed have a direct bearing on the achievement of net impacts, whether they be on earnings or welfare dependency.

The fact that performance standards can play a critical role in a program's long-term success is illustrated in Exhibit 5.3. Here, we present the correlations

---

[1]As discussed earlier, part of this result likely reflects the use of background characteristics to estimate individual trainee impacts.

EXHIBIT 5.2

CORRELATIONS BETWEEN SELECTED BACKGROUND CHARACTERISTICS

AND VALIDATION BENCHMARKS

Selected Background Characteristics	VALIDATION BENCHMARKS	
	EARNINGS GAINS	PUBLIC WELFARE DEPENDENCY
Pre-Program Work History		
Ever Worked for Pay	-0.081[a]	0.009
Longest Pre-Program Job (in weeks)	-0.096[a]	0.039
Months of Continuous Pre-Prog. Unemployment	0.128[a]	-0.123[a]
Have License and Car	-0.043[c]	0.071[a]
Pre-Program Welfare Dependency		
Months Since First Received AFDC	0.082[a]	-0.094[a]
Average Monthly Public Benefit Payment (prior 12 months)	0.049[c]	-0.154[a]

[a] Significant at 99%
[b] Significant at 95%
[c] Signficiant at 90%

EXHIBIT 5.3

CORRELATIONS BETWEEN UNADJUSTED CANDIDATE PERFORMANCE MEASURES AND SELECTED BACKGROUND CHARACTERISTICS

Unadjusted Candidate Measures	Background Characteristics					
	Ever Worked For Pay	Longest Pre-Program Job (in weeks)	Months of Continuous Pre-Program Unempl.	Have License and Car	Months Since First Received AFDC	Average Monthly Public Benefit Payment (prior 12 mos)
Placed in Job	0.042[c]	0.058[b]	-0.024	0.097[a]	-0.004	-0.012
Average Placement Wage	0.061[b]	0.087[a]	-0.042[c]	0.103[a]	0.011	0.035
Employment Status						
3 months	0.065[b]	0.068[a]	-0.063[b]	0.077[a]	0.014	-0.005
6 months	0.034	0.077[a]	-0.067[a]	0.067[a]	0.029	0.005
Earnings Status						
3 months	0.078[a]	0.083[a]	-0.083[a]	0.106[a]	0.003	0.011
6 months	0.074[a]	0.086[a]	-0.089[a]	0.106[a]	0.002	0.016
Welfare Status						
3 months	-0.067[a]	-0.031	0.050[c]	-0.083[a]	0.037	0.090[a]
6 months	-0.096[a]	-0.016	0.038	-0.075[a]	0.025	0.058[b]
Welfare Benefits Level						
3 months	-0.047[c]	-0.069[a]	0.047[c]	-0.121[a]	0.026	0.344[a]
6 months	-0.054[b]	-0.066[a]	0.056[b]	-0.131[a]	0.030	0.345[a]

[a] Significant at 99%
[b] Signficiant at 95%
[c] Significant at 90%

between the same six background characteristics and each of the candidate measures. As can be seen, of the sixty separate correlations tested, only twenty are insignificant; and of these, all but four are associated with the two characteristics reflecting prior welfare dependency. These characteristics are not well correlated with the employment and earnings indicators, but they are with those reflecting one's welfare status and benefit level.

As a final check on the results, we tested for the best overall indicator and the two best indicators of each validation benchmark, using all of the (unadjusted) candidate performance measures and the background characteristics. The results are shown in Exhibit 5.4. Consistent with the earlier findings, background characteristics are selected as among the top two predictors of both validation benchmarks. In fact, the average monthly benefit level in the year prior to program enrollment was selected as the best predictor of net reductions in welfare dependency. We can thus say that factors affecting one's potential employability play a key role in the production of net program impacts on both earnings and welfare dependency. We can also see that these same factors influence performance achievement. To the extent that targeting and program mix decisions are designed to maximize the chances of achieving established performance standards, it is clear that such decisions will run counter to the goals of increasing earnings and reducing welfare dependency.

As discussed above, the intuition behind these findings has been incorporated into JTPA's performance management system through the optional adjustment model. However, there are reasons to believe that the model approach alone, will not be sufficient to stimulate services to the AFDC population and especially those that are hard-to-employ. Above and beyond reasons related to the technical merits of the model, a key consideration is that in a practical planning sense, its use is strictly voluntary. In other words, an SDA must decide first to target on AFDC recipients before it can take advantage of the adjustments provided by the model. Thus, we can see that availability of modeled adjustments alone will not necessarily stimulate service delivery to the AFDC population.

Compounding this is the problem faced by service vendors. It is at this level that actual enrollment decisions are made, and where creaming can be most evident. Since the adjustments currently offered to SDAs are not directly passed on to vendors, contractors operate largely with unadjusted performance scores. To the extent that they are motivated to achieve high levels of performance through performance-based contracts, for example, they have a clear incentive to cream from among the eligible population.

EXHIBIT 5.4

CORRELATIONS BETWEEN UNADJUSTED CANDIDATE MEASURES, SELECTED BACKGROUND CHARACTERISTICS AND ALTERNATIVE VALIDATION BENCHMARKS

	Earnings Gains					Public Welfare Dependency				
Pre-Program Work History	Best Pre-Prog. Work History Ind.	Best Pre-Prog. Welfare History Indicator	Best Pre-Program Indicator	Best Overall Indicator	Two Best Overall Indicators	Best Pre-Prog. Work History Ind.	Best Pre-Prog. Welfare History Indicator	Best Pre-Program Indicator	Best Overall Indicator	Two Best Overall Indicators
Ever Worked for Pay										
Longest Pre-Program Job (in weeks)										
Months of Continuous Unemployment	0.464[a]		0.464[a]		.314[a]	-0.538[a]				
Have License and Car										
Pre-Program Welfare Dependency										
Average Monthly AFDC Payment										
Average Monthly Total Benefit Level							-0.181[a]	-0.181[a]	-0.181[a]	-0.241[a]
Months Since First Received AFDC		0.184[a]								
Unadjusted Candidate Measures										
Placed in Job										
Average Placement Wage										
Employment Status										
3 months										
6 months										
Earning Status										
3 months										
6 months				0.258[a]	0.280[a]					

EXHIBIT 5.4 (continued)

COMPARISON OF CORRELATIONS BETWEEN ALTERNATIVE VALIDATION BENCHMARKS

	Best Pre-Prog. Work History Ind.	Best Pre-Prog. Welfare History Indicator	Best Pre-Program Indicator	Best Overall Indicator	Two Best Overall Indicators	Best Pre-Prog. Work History Ind.	Best Pre-Prog. Welfare History Indicator	Best Pre-Program Indicator	Best Overall Indicator	Two Best Overall Indicators
Welfare Status										
3 months										
6 months										
Welfare Status										
3 months										
6 months										
Welfare Benefits Level										
3 months										
6 months										0.025[a]
Multiple Correlation Coefficient ($\sqrt{R^2}$)	0.128[a]	0.082[a]	0.128[a]	0.142[a]	0.200[a]	0.123[a]	0.195[a]	0.195[a]	0.195[a]	0.242[a]

Note: Indicator entries reflect parameter estimates from regression equations

[a] Significant at 99%
[b] Significant at 95%
[c] Significant at 90%

But the presence of such incentives do not necessarily mean that creaming occurs or that performance standards are the most critical factor in shaping the targeting and enrollment decisions of SDAs. As we discussed earlier, several other factors bear upon the choice of program services and of individuals to enroll in programs. Only by better understanding the full context in which such decisions are made is it possible to determine the expected impact of improvements in performance standards, and what such improvements should be. We now turn to this issue by presenting the key findings that emerged from our telephone interviews and visits to the field.

5.2 Performance Standards, Targeting, and Program Mix: Findings from the Field

In Chapter 2 of this report, we presented and discussed the major issues that state and local practitioners face in providing services to the AFDC population, and in targeting resources to those deemed hard-to-employ. Of all the issues raised and discussed, three, in particular, stand out. The first centers on both the importance of and difficulty with achieving inter-agency coordination. At both the state and local levels, the need for coordination between JTPA and welfare as well as other program agencies such as education and social services, were consistently cited as key to effectively serving the welfare population. At the same time, our respondents also pointed to the extraordinary difficulty involved with achieving a coordinated response to the employment, training, and educational needs of this population group.

At the state level, turf battles, ambiguity over respective roles and responsibilities, fundamental differences in perspective over the appropriate mix between work and welfare, and lack of clarity over whom to target prevented much of the coordination viewed as necessary from materializing into a routine pattern. As a result, it often took top level gubernatorial support to pull the relevant agencies together, and special demonstrations or welfare reform initiatives to form a coordinated program response.

At the local level, even in the presence of state-wide demonstrations, coordination remained a moving and evasive target. The three SDAs we visited each indicated the extreme difficulty they had experienced in attempting to establish routine inter-agency referrals, and mechanisms for information sharing and on-going support. The pressing priorities of each agency and the limited incentives they had to adopt full-fledged coordination curtailed the ability of localities to join forces and represent a united program front.

The second key issue raised was the difficulty encountered in providing comprehensive and flexible services. Most of our respondents agreed that it was inappropriate to prescribe a fixed service mix for the AFDC population. Rather, most viewed it as necessary to offer a flexible service mix that could respond to a varying need for education, training, and support services, including child care, health benefits, and when needed, housing assistance. While limited coordination was frequently cited as contributing to this problem, so were limited resources, federal regulations governing the availability of support services and public housing, and the absence of guaranteed health insurance.

Finally, compounding these two issues were the disincentives many welfare recipients face in choosing work over welfare. In light of the low quality jobs available to many welfare recipients and the loss of key benefits, several of our respondents pointed to the dilemma faced by those on welfare.[1] This is well illustrated in Exhibit 5.5, which shows for the state of Florida how a mother of two will fare in terms of net additional income gained from leaving welfare and going to work full-time. As can be seen, the net addition is a mere $.02 an hour. Although this transition would result in over $500 monthly net savings to the public, without guaranteed continued health care, child care assistance and other supportive services, welfare recipients were characterized as seeing this as a bad bargain.

To a large extent, the comments made by our respondents reflected the overwhelming nature of the constraints they face in attempting to provide more and better services to the welfare populations. By all accounts, state and local respondents shared a general commitment to improving service delivery to the AFDC population, but believed that a number of basic institutional constraints prevented them from doing so as effectively as they would like.

Another important possible constraint is that of performance standards. Given the controversy surrounding the perverse incentives that performance standards may introduce, we also queried our respondents about whether performance achievement further constrained their ability to serve the welfare population. While the general

---

[1]Recall that during the initial three months of employment, recipient benefit levels are reduced by roughly $.33 for each dollar earned. After that point, recipients experience a near dollar reduction in benefits for each dollar earned. In addition, the continuation of child care are health benefits after leaving AFDC (partially or fully) vary by state. Across all states, however, both benefits end within a six to nine month period.

EXHIBIT 5.5

PUBLIC ASSISTANCE VS. WORK COMPARISON OF NET INCOME

FOR AFDC RECIPIENTS IN THE STATE OF FLORIDA

Example: Mother with a son (age 7) and a daughter (age 4).

Public Assistance

AFDC	$264.00/month	
Food Stamps	164.00/month	
Less Medical Coverage (medicaid)	nc	
Less Public Housing Rent	55.00/month	
Net Monthly Income	$373.00/month	

Work (40 hours per week at $5.00 per hour)

Gross Wage Income	$866.00/month	
AFDC	14.00/month	
Food Stamps	37.00/month	
Less Federal Income Tax	69.00/month	
Less F.I.C.A.	62.00/month	
Less Public Housing Rent	125.00/month	
Less Average Child Care	180.00/month	
Less Transportation	25.00/month	(est.)
Less Cost of Work Clothes	85.00/month	(est.)
Net Monthly Income	$376.00/month	(est.)

Income Gain/(Loss) From Working

Net Working Income	$376.00/month	(est.)
Net Public Assistance Income	373.00/month	
Income Difference	$ 3.00/month	(est.)
Net Earnings Per Hour of Work	$ .02/hour	(est.)

response was affirmative, most of those interviewed did not view the standards as the key constraining factor. Key details regarding respondent perceptions of the role of performance standards are discussed below.

We began the interviewing process with a series of questions designed to identify the role that performance played in program and target group planning. We were particularly interested in the extent to which program plans were driven largely by the desire to meet standards set by the state. By and large, we found that SDAs and administrators of welfare-sponsored employment and training programs did not, first and foremost, plan their program and client mix around performance standards. While performance achievement appeared to be a more significant goal among JTPA staff, they, like their welfare counterparts, planned programs in response to other factors, such as community pressures and politics, perceived client needs, cost constraints, and precedent established by the previous mix of programs. In no case were we told that performance achievement was a critical and central force in the development of targeting criteria and program budgeting plans.

This is not to say that performance achievement was ignored. Indeed, the SDAs were quite mindful of the impact of their program mix on their ability to meet their performance standards; and in many cases, they avoided taking large program risks so as not to place performance achievement in jeopardy. But for the most part, SDAs and welfare offices did not make large changes in their programs or targeting priorities in response to their perceived ability to meet expected levels of performance.

Part of the reason for this is that the SDAs we visited did not provide services directly. They planned, administered, and funded programs, and in so doing, passed on the tension between performance achievement and targeting to their program operators. Indeed, it was at the actual service delivery level where conflicts between performance achievement, client enrollment, and program investments were observed.

Central to the conflict faced by program operators are performance-based contracts. These contracts establish specific unit-costs for individual performance benchmarks, such as enrollment levels, participation rates, and program outcomes, including placement and retention. For the SDAs, these contracts are viewed as an efficient tool for ensuring accountability and a focus on performance outcomes. For contractors, however, these contracts represent an important source of conflict when viewed in the context of matching services to the needs of AFDC recipients, including those that are hard-to-employ.

As an illustration, consider the benchmarks used by one of the SDAs we visited. In this case, 45 percent of the total unit cost of serving an individual is contingent upon the two program outcomes, placement and retention on the job for at least 30 days. Since nearly one-half of the cost of serving each individual is tied to events that occur at and after termination, contractors in this SDA have a clear incentive to serve individuals who have a strong likelihood of obtaining and staying on a job.

The SDAs we visited recognized this conflict and have made attempts to limit it primarily by negotiating total unit costs that best reflect the target groups receiving priority. However, the SDAs also displayed some degree of reluctance to negotiate too much away, or in other words, lower expected outcomes and raise costs limits too high. This reluctance stemmed from a growing perception among PIC members and SDA staff that performance achievement was critical, on an intuitive as well as technical level. As discussed in our interim report, the emphasis of JTPA on performance achievement has created a culture in which the standards have become quite prominent in the minds and management systems of program administrators, overriding other key management concerns and the primary importance of focusing on individual employability development.

Yet, we heard little in the way of suggestions that JTPA's performance standards should be dropped or even substantially changed. While most of our respondents believed that the standards setting process could be improved, they viewed the chief challenge in providing better services to AFDC recipients in terms of addressing more basic institutional and program constraints. If changes to the performance standards are to be made, then SDAs pointed to three issues that should receive careful attention. These include:

1. elimination of the cost standards;

2. improvements to the adjustment model that are designed to better capture key variables that affect local performance; and

3. larger financial incentives for performance achievement, in general or with respect to services provided to a selected target group, such as AFDC recipients.

These findings from the field suggest two important implications. First, they indicate that performance standards operate in an environment already constrained by institutional factors largely out of the control of any one agency or program serving the AFDC population. Limited inter-agency coordination, support services, and remedial

education as well as strong work disincentives make the task of serving the welfare population difficult. Unless changes are made to ameliorate these limitations, the delivery of services to the AFDC population will continue to be hampered.

Second, and within this context, there is evidence to suggest that performance achievement may well be contributing to the difficulty in serving AFDC recipients. While our results do not indicate that the standards are among the most important limiting factors, they do seem to play a role in shaping program and targeting decisions. Especially at the local contractor level, we observed a tension that many have argued exists between performance standards and targeting. Thus, while it appears that the elimination of this tension will not necessarily unleash a plethora of services to the welfare population, improvements to the performance management system can help to make incremental improvements. We now turn to this issue of creating a performance standards system that is more supportive of serving AFDC recipients.

5.3 Forging a Responsive Performance Management System

We began this report with a discussion of the inconsistencies that exist between the basic structure of JTPA and the growing expectation that it serve AFDC recipients who are are hard-to-employ. As we showed, such inconsistencies stem from the Act's broad targeting criteria, emphasis on state and local decision-making, and limited emphasis (at least from a cost perspective) for support services. We also presented the concerns raised by program practitioners at the state and local levels over many of these same types of issues. By all accounts, the evidence we have observed suggests that efforts to target on AFDC recipients and to provide them with effective services operate within an environment constrained by many institutional factors which are largely out of the direct control of any single state or local agency.

When considering the role that a responsive performance management system can play, it is thus important to place it in this context. While we have also observed that standards can have an effect on enrollment and program investment decisions, it appears that they are not the chief binding constraint to serving the welfare population. But it does appear that basic improvements can be made that should be expected to better stimulate service delivery to AFDC recipients.

To begin, the results from our validation tests suggest that the use of many different performance measures may not be superior to relying on just a few. We found that, on the basis of statistical validity, earnings and welfare status or benefit levels during the post-program period are the best predictors of both longer-term earnings gains

and reductions in welfare dependency. While the six month indicators outperformed those measured at three months, the observed differences were not always large.

In making a final selection of performance indicators, these statistical results need to be considered along with practical considerations involved with obtaining the data required to measure a given indicator, and using a specific indicator in a timely fashion. For example, while the six month measures were found to have greater predictive validity than the three month ones, it is not clear that the difference warrants the additional time required to obtain data and measure the indicator. Similarly, the superiority of welfare benefit levels over welfare status as a potential indicator must be balanced with the increased complexity involved with obtaining the data needed to measure it. Ultimately, the final choice of candidate performance measures must take these factors into account along with the specific validation findings. Costs, ease of data collection and measurement, and timely use of information must each be carefully weighed in order to develop a technically sound performance management system that has practical management appeal and usefulness.

Our results have also indicated that there is a strong correlation between the candidate performance measures and background characteristics that reflect employment and earnings potential. This implies that the potential exists to encourage service providers to enroll individuals with characteristics that will maximize their chances of achieving high performance scores.

One way to address this is through statistical modeling. Indeed, the intuition underlying the Department of Labor's adjustment model is that by removing the effects of client characteristics on performance achievement, SDAs (and program operators) will have limited incentives to engage in selective enrollment.

Whether or not this approach is effective has come under increasing scrutiny. The states and SDAs we talked with indicated that the available adjustment models have only a limited influence on targeting and program mix decisions. As discussed earlier, the small size of the adjustment factors and the relative ease by which SDAs can meet their performance standards contribute to this perception. However, beyond reasons related to the technical merits of the model, an additional consideration is that using the model, in a practical planning sense, is strictly voluntary. In other words, an SDA must decide first to target on AFDC recipients before it can take advantage of the adjustments provided by the model.

Compounding this is the problem faced by service vendors. It is at this level

that actual enrollment decisions are made, and where creaming pressures were most evident. Since the adjustments currently offered to SDAs are not directly passed onto vendors, contractors operate largely with unadjusted performance scores. To the extent that they are motivated to achieve high levels of performance through, for example, performance-based contracts, they have an incentive to cream from among the eligible population.

Thus, the availability of statistical adjustment models alone will not necessarily stimulate service delivery to the AFDC population or sub-groups of it. Moreover, while it is possible to improve the modeling process through the inclusion of additional variables, such efforts should not be expected to produce large changes in current behavior.[1]

There are, however, certain steps that can be pursued in conjunction with statistical modeling. One would be to increase the incentives service providers have in serving AFDC recipients. Most of the state and SDA respondents we spoke with viewed JTPA's six percent incentive funds as a viable mechanism to induce certain types of targeting and program mix decisions. However, they also cited the need to better direct the use of these funds to targeting issues and to also increase the funding level allocated to incentive dollars.

The second would be to make the statistical adjustment models available to local vendors. Currently, this is not the case. While SDAs rely on the models, they do not and cannot make them readily available to their vendors during the planning process. As discussed earlier, this leaves program vendors without the benefit of the adjustment models, limited as they may be.

Vendor use of the adjustment models could be accommodated in one of two ways. On the one hand, SDAs could include the adjustment factors from the existing models in vendor RFPs as an optional (or required) planning and bidding tool. On the other hand, locally-based models, using SDAs micro-data could be developed and used in the same fashion.

A third and final step would be to increase the level of technical assistance provided to states and particularly SDAs regarding the provision of effective services to

---

[1]See Appendix 4-A for a presentation of key variables affecting the candidate performance measures.

the AFDC population. Our results indicate, for example, that assistance in the area of identifying the hard-to-employ population would help JTPA to more readily target its resources and distinguish among the various types of (AFDC) participants it potentially could serve. Similarly, technical assistance in developing appropriate program designs and coordinating with other agencies to leverage resources would also be quite helpful, given the difficulties we observed in the area of inter-agency coordination.

Taken together, these types of changes to the current performance standards structure should be expected to provide some relief from existing creaming incentives. However, as discussed earlier, absent selected changes in the targeting and program provisions of JTPA, they will not necessarily produce large changes in state and SDA choices over who to serve and in what programs.

If there is an interest in using the performance standards system to stimulate greater service levels to AFDC recipients, it would appear that some other, or additional approach may be warranted. On the basis of our validation research, one viable option would be to set target group performance standards. Short of amending the targeting provisions of the JTPA legislation, target-group performance measures would exert a clear pressure on the state and local levels to provide services to selected population sub-groups.

There are two ways in which target-group performance measures could be established. One would be to set service shares for pre-specified groups. Since JTPA already contains service share requirements for the AFDC population, the use of such performance measures would have to focus on sub-groups of the welfare population, such as long-term recipients, teen parents, or those with very limited work histories. In view of the wide range of definitions currently in use by the states to identify hard-to-employ AFDC recipients, such efforts would likely have to consider a range of options and result in the selection of one that best meets the interests of the states and their SDAs.

Another option for setting target-group performance measures would be to establish specific outcomes for either the AFDC populations in total, or for selected sub-groups of it. In this case, the focus would be placed on achieving pre-established levels of performance, in contrast to services shares. The advantage of this option is that it maintains JTPA's focus on outputs, as distinct from process.

But the use of service share measures would not have to replace the current emphasis on program outputs. Rather, they could be viewed as a first-order condition of performance. In other words, employment and training programs would be required first

to achieve certain levels of service to pre-specified target groups. Having achieved this, they would then be judged on the basis of how well they performed in terms of program output measures. Thus, modeling would continue, but it would not be viewed as the only vehicle for encouraging services to particular groups that would otherwise be avoided due to concerns over performance achievement.

Adopting such an approach also raises certain issues. For one thing, creating target group performance measures could be viewed as running counter the basic principles of decentralization underlying the JTPA program. Second, the performance management system would have to be provided with the support necessary to help overcome existing barriers to coordination, and identify as well as implement the most effective way to organize and provide services to special population groups. Absent these efforts or new legislation to support them, our findings indicated that even the best performance management system will not alter the behavioral preferences of program operators.

Ultimately, the choice over how to structure a performance management system for targeted welfare programs will depend on the objectives established for it. The results of this study have indicated that institutional factors and legislative parameters largely constrain the ability of states and localities to dedicate resources to the needs of the AFDC population and sub-groups of it that are hard-to-employ. Absent efforts to resolve these problems, the performance standards can be used to achieve specific targeting objectives, but not without measures that may well be viewed as counter to the basic principles of the existing employment and training system. Unless such steps are taken, it is clear that there is a viable role for performance standards, but one that should not be viewed as necessarily resolving large, more basic issues.

REFERENCES

Bell, Stephen H., Evaluation of the AFDC Homemaker-Home Health Aide Demonstrations: Trainee Public Program Benefits. Washington, DC: Abt Associates Inc., December 1987.

Bell, Stephen.; Burstein, Nancy L.; and Orr, Larry L., Evaluation of the AFDC Homemaker-Home Health Aide Demonstrations: Overview of Evaluation Results. Washington, DC: Abt Associates Inc., December 1987.

Bloom, Howard S., "Accounting for No-Shows in Experimental Evaluation Designs." Evaluation Review, April 1984: 225-246.

Enns, John H.; Bell, Stephen H.; and Flanagan, Kathleen L., Evaluation of the AFDC Homemaker-Home Health Aide Demonstrations: Trainee Employment and Earnings. Washington, DC: Abt Associates Inc., December 1987.